AERIALS FOR METRE AND DECIMETRE WAVE-LENGTHS

AERIALS FOR METRE AND DECIMETRE WAVE-LENGTHS

BY

R. A. SMITH

M.A., PH.D., A.M.I.E.E.

*Superintendent, Physics Department,
Telecommunications Research Establishment,
Ministry of Supply*

CAMBRIDGE
AT THE UNIVERSITY PRESS
1949

PREFACE

The extension of radio technique to extremely short wave-lengths has increased enormously the scope of aerial design. Half-wave and full-wave dipoles have reasonable dimensions when the wavelength is less than a few metres, and arrays for a multitude of special purposes may be designed. The selection of the range of wave-lengths for which aerials are discussed in this volume has been somewhat arbitrary. The upper limit has been set at about 12 m., determined mainly by two considerations. First, for greater wave-lengths rigid dipoles become cumbersome and are not much used. Secondly, 12 m. is the longest wave-length used for radar, the widespread use of which at shorter and shorter wave-lengths has led to a great development in aerial technique.

By the microwave region is usually meant that extending from 10 cm. to shorter wave-lengths, so that the lower limit has been taken as 10 cm. Many new developments in aerial technique have taken place in the microwave region, and a separate volume of this series will be devoted to the subject of microwave aerials.

The 'metre' band has therefore been taken arbitrarily as extending from 12 to 1 m., and it is with this band which we shall mainly be concerned. A short chapter is, however, included on aerials for the 'decimetre' band extending from 1 m. to 10 cm. Only two parts of this band, which have been extensively developed, will be discussed, that lying near 50 cm. and that lying near 30 cm.

This volume does not attempt a comprehensive description of the great variety of aerial systems used for short-wave radio applications. Indeed, most radio applications have aerial systems with their own characteristic features, so that the number of different arrangements which have been used is enormous. All that has been possible is the selection of a few applications as illustrating general principles. These illustrations have largely, though not entirely, been taken from radar applications, since it is here that perhaps the most outstanding developments in aerial design have taken place.

To my colleagues at the Telecommunications Research Establishment I should like to express my gratitude for countless discussions;

in particular to Mr F. E. Lutkin and Mr R. H. J. Cary, who have worked with me on a number of aerial problems and whose wide experience has always been available; also to Dr Denis Taylor, Mr G. E. Bacon and others whose help with illustrations from radar in the $1\frac{1}{2}$ m. and 50 cm. bands has been invaluable. I should also like to express my gratitude to Dr H. G. Booker and Dr G. G. Macfarlane for many valuable discussions on the theoretical aspects of aerials. Mr J. A. Ratcliffe read the book in manuscript and I am indebted to him for a number of valuable suggestions.

Finally, I wish to thank the Ministry of Supply for permission to publish this volume and for the use of a number of drawings and photographs. Most of such knowledge and experience of the subject as I possess has been gathered while engaged in the wartime radar research programme at the Telecommunications Research Establishment.

<div align="right">R. A. SMITH</div>

Telecommunications Research Establishment
Ministry of Supply

21 *March* 1947

CONTENTS

Chapter 1. RESONANT DIPOLES

Chapter 2. AERIAL ARRAYS

Chapter 3. ELEMENTARY THEORY OF AERIALS

LIST OF PLATES

Chapter 1

RESONANT DIPOLES

1·0. Introduction

Until the advent of radar very little research had been carried out on aerials for wave-lengths shorter than a few metres. A great deal was known, however, about the properties of aerials for longer wave-lengths, and the need for a new outlook on aerial design was being realized as a result of problems arising from short-wave communications and television. Because of a legacy from broadcasting there was at the beginning of the war a widespread belief that although considerable care must be taken with design of transmitting aerials, any old piece of wire would do for a receiving aerial. This belief was very hard to break down, and as a consequence research on receiving aerials was neglected. For long wave-lengths the reason for the belief is well founded. At wave-lengths greater than about 20 m. receivers are readily available which generate less noise than is picked up by even a quite inefficient aerial. The background noise (which we shall discuss in more detail later) determines the limit of perceptible signals, and improvement of the aerial increases both signal and noise in the same ratio so that no advantage is gained by having an efficient receiving aerial.

For shorter wave-lengths, however, the situation is quite different, and the limit of sensitivity is set by noise generated in the receiver itself. Any improvement in aerial performance then gives an increase in the ratio of signal to noise and is therefore worth while. For short wave-lengths, therefore, it is essential that receiving aerials as well as transmitting aerials should work at maximum efficiency.

In many applications a receiving aerial is simply a means of picking up the required radio energy and conveying it to the receiver. For radar, however, as used to locate the position of a target the receiving aerial is the part of the instrument which makes the measurement of azimuth and elevation and so determines the accuracy of the system. The need for care in design is then obvious.

The development of television had already brought to light the need for accurate matching of both transmitting and receiving aerials so as to avoid spurious reflexions which, even with the very

short delays normally involved in the passage of the signal from aerial to receiver or transmitter, may still be of significance. In radar the same considerations apply, so that accurate matching is an essential feature of radar aerials.

In the design of transmitting aerials for long waves the chief problem is to minimize the losses in the earth. If the aerial consists, as it frequently does, of a vertical wire with one end near the ground, a length of approximately a quarter of a wave-length is required before the system is resonant. For shorter lengths the aerial is highly reactive. For short waves, however, use can be made of half-wave or full-wave resonant dipoles. Moreover, they can readily be raised a few wave-lengths above the ground and highly efficient aerials made, in which the losses are but a few per cent, practically all the power from the transmitter being radiated. Moreover, vertical or horizontal polarization can be used at will.

Such dipoles may be fed by a balanced transmission line, or by a coaxial line with balance-to-unbalance transformer. This technique was quite well known before the war[1] and will only be mentioned where advances of note have been made. In what follows it will be assumed that the elementary principles of transmission along lines are known, and that the reader is familiar with the method of matching a load to a transmission line by means of a stub or transformer.

1·1. The resonant dipole

Since the resonant dipole forms the basis of most short-wave aerials, we shall briefly summarize some of its more important properties. Such a dipole is shown in fig. 1 fed by a balanced transmission line. The dipole consists of two pieces of straight tubing or wire AB, CD, each of length l with spacing S at the centre. It is well known that when l is approximately equal to an integral number of quarter wave-lengths the impedance as seen at AC is resistive and the dipole behaves like a resonant or anti-resonant circuit. For other values of l the impedance of the dipole is complex and its form is shown in fig. 2. The values of l most commonly used are approximately $\frac{1}{4}\lambda$ and $\frac{1}{2}\lambda$. In the first case we have a half-wave dipole whose centre impedance at resonance is of the order of 70 ohms for a wide range of thickness of tubing. When $l = \frac{1}{2}\lambda$ we have a full-wave dipole, and in this case the centre impedance

depends markedly on the diameter of the tubing as shown by table 1. The values given in this table have been taken from a theoretical paper by Schelkunoff(2), but are in good agreement with experimental values. Experiments on half-wave dipoles(3) have shown that the resonant resistance varies with the centre spacing,
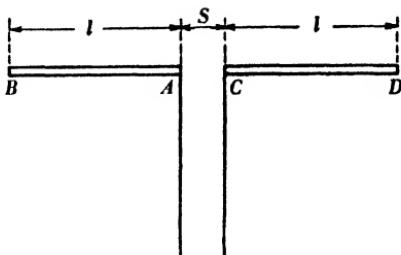
Fig. 1. Dipole fed by balanced transmission line.
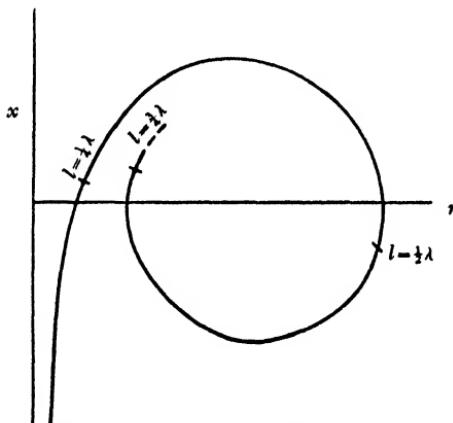
Fig. 2. Impedance of dipole, $r + jx$.

but it is not known to what extent the effect is found for full-wave dipoles. The values of l near $\frac{1}{2}\lambda$ for anti-resonance are also shown in table 1.

Although a large amount of theoretical and experimental work has been applied to the determination of the impedance of the half-wave dipole the results are far from satisfactory, in that theory and experiment do not agree well and experiments are mutually conflicting. This has been mainly due to differences in feeding arrangements. We shall examine the theoretical position in Chapter 4. Experimental determinations have been scanty and have been mainly

carried out in the course of design of particular aerial systems and are limited in scope. Such measurements have been reported in a paper by R. A. Smith and C. Holt Smith dealing with receiving aerials for C.H. radar stations(3), and are shown in fig. 3, together

Table 1. *Full-wave dipole*

Wave-length/ diameter $\lambda/2a$	Anti-resonant resistance in ohms	Value of $2l/\lambda$ for anti-resonance
100	900	0·870
200	1,300	0·896
400	1,700	0·916
1,000	2,400	0·937
2,000	3,000	0·945
4,000	3,600	0·951
10,000	4,600	0·958
100,000	8,000	0·967

with some unpublished measurements made by W. Cochrane for the same purpose. The main fact which emerged from these measurements* was that the resistive part of the impedance depends on the spacing at the centre of the dipole and on the method of connecting the transmission lines. For dipoles made of tubing it is usually lower than the frequently quoted figure, 73 ohms, which is the theoretical value for an infinitely thin dipole. Measurements made at the National Physical Laboratory in 1941(4) indicated a resistive component higher than 73 ohms for a number of dipoles, but these were made with a relatively large centre spacing, and as will be seen from fig. 3 the resistive part of the impedance increases with spacing S.

For the measurements recorded in fig. 3 the dipole was fed by open-wire transmission lines with various spacings. In fig. 4 is shown the measured impedance of a dipole fed by bonded concentric lines(3). It should be noted that the form of the bond was found to have an appreciable effect on the impedance. The values shown in the figure were obtained using a bond of very low impedance. Values given by Schelkunoff(2) for the resonant length of a half-wave cylindrical dipole are given in table 2.

The value obtained for $2l/\lambda$ from fig. 3 for a dipole in which the ratio wave-length/diameter = 200 is 0·468 which is only slightly

* These were carried out by a number of research workers in addition to the authors of the paper referred to, including Dr B. B. Kinsey and Mr G. H. Palmer.
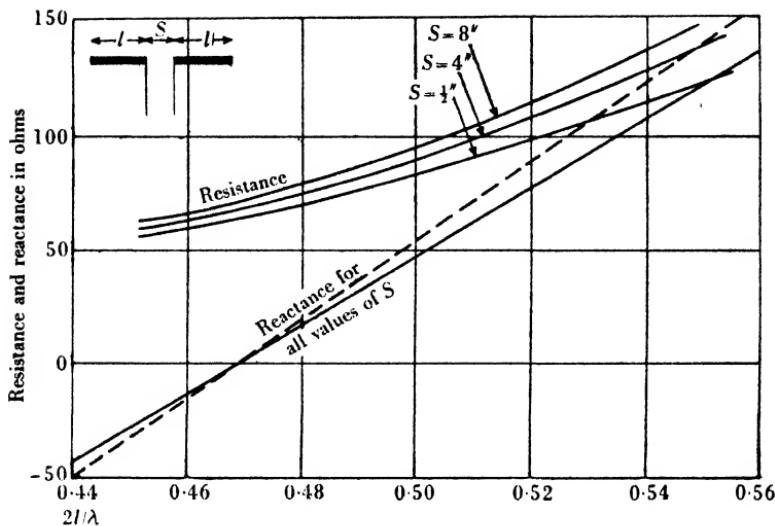
Fig. 3. Impedance of dipole with various values of centre spacing.
———— Smith and Holt Smith ($\lambda = 6$ m. $a = 0.625''$ $\lambda/4a = 100$)
– – – – Cochrane   ($\lambda = 7.2$ m. $a = 0.375''$ $\lambda/4a = 200$)
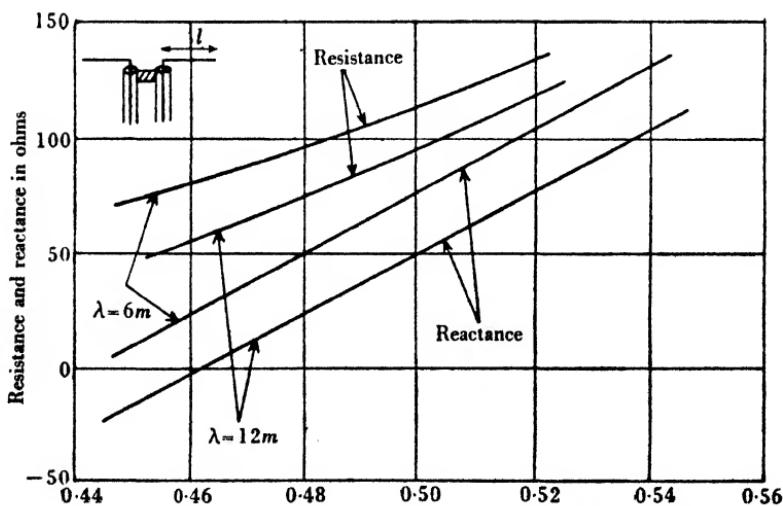
Fig. 4. Impedance of dipole fed by bonded concentric lines.

smaller than the value, 0·471, obtained theoretically by Schelkunoff. It will be noticed that the value was found to be independent of the spacing S (see also fig. 19, Chapter 4).

For values of l less than the resonant value the reactive part of the impedance is negative, and for larger values of l it is positive till the next resonance is reached near $l \simeq \frac{1}{2}\lambda$, when it passes through zero and again becomes negative as shown in fig. 2. This process is repeated for each subsequent resonance.

Table 2. *Half-wave dipole*

Wave-length/ diameter	Value of $2l/\lambda$ for resonance	Resonant resistance in ohms
200	0·471	61·6
400	0·475	63·6
1,000	0·479	65·3
4,000	0·484	67·2
10,000	0·486	68·1
100,000	0·489	69·2

The theoretical values for the resonant resistance of a dipole are also given in table 2. In practice, however, these values are seldom found to be in agreement with experiment, as the resistance depends on the width of the gap as shown in fig. 3. In view of the lack of consistent measurements covering a wide range of dipoles a designer who needs a very accurate estimate of the impedance of a half-wave dipole must measure it under the conditions of use. In any case, as we shall see, supporting insulators have a marked effect on the impedance.

Although the resistive part of the impedance does not vary much with the ratio of diameter to wave-length the rate of change of reactance with length or frequency near the resonant point certainly does, being greatest for thin dipoles. For all ratios of wave-length λ to diameter $2a$ greater than about 200 the input resistance of a half-wave dipole increases at a rate of about 2·5 ohms for 1 % increase in length or in frequency. The reactance, however, increases at a rate of $\pi Z_0/200$ ohms for a 1 % increase in length or in frequency, where Z_0 is equal to 120 ($\log_e 2l/a - 1$). The slopes of the curves in fig. 3 agree almost exactly with these values (see Chapter 4).

The main advantage of using a dipole at a resonant or anti-resonant point is that, since it is purely resistive, we can arrange to

have either a current or voltage minimum at the feed-point, and this is frequently convenient. If we happen to have available a transmission line with characteristic impedance equal to the input resistance it will match the dipole without need for further elaboration. If it is not equal, the dipole may be matched, when resistive, by means of a single quarter-wave transformer.

1·2. Effect of parasitic dipole on half-wave dipole

When a free dipole, in the form of a continuous rod of length nearly equal to $\frac{1}{2}\lambda$, is brought near a half-wave dipole it has a marked effect on its centre impedance and on its polar diagram. A theoretical discussion of this effect will be given in Chapter 5. If the parasitic dipole (as the continuous rod is called) has a length greater than half a wave-length it acts as a reflector, and if its length is somewhat less than half a wave-length it acts as a director. As a reflector the parasite decreases considerably the radiation towards it and augments that in the opposite direction. As a director it increases the radiation towards it and reduces that in the opposite direction. A detailed treatment will be given in Chapter 5 in which the effect of varying the length of the parasite will be discussed.

1·3. The folded dipole

The presence of a parasitic element near a half-wave dipole has the effect of lowering considerably the resistive part of its impedance.
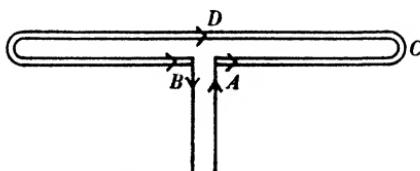
Fig. 5. Folded dipole.

When the dipole has a free-space resistance of 70 ohms the resulting low value is often inconveniently small and it is desirable to start with a half-wave aerial whose centre impedance in free space is considerably greater than 70 ohms. A 'folded dipole' is a convenient aerial of this type. The form of this aerial is shown in fig. 5. The over-all length is usually slightly less than $\frac{1}{2}\lambda$ in order to obtain a resistive feed, but depends on the thickness of the material from which the dipole is constructed. The centre impedance of a folded

dipole is about four times that of a half-wave dipole. This may be seen roughly as follows. Regarding the folded aerial wire as a transmission line, the point D will be electrically half a wave-length away from A and will have current equal to that at A but 180° out of phase. Due to the wire being bent round, the current will flow in the same direction at D as at A. We have therefore effectively doubled the radiating current at the centre of the aerial. Let R_0 be the resonant resistance of a half-wave dipole and R_d that of the folded dipole. Then since these produce almost identical polar diagrams the power radiated by the folded dipole will be approximately equal to that from a half-wave dipole carrying twice the current. Hence we have

$$R_0(2i)^2 = i^2 R_d$$

or
$$R_d = 4R_0. \tag{1}$$

R_d has therefore a value of the order of 300 ohms.

REFERENCES

(1) Ladner, A. W. and Stoner, C. R. *Short-wave Wireless Communication*, 4th ed. (1942). Chapman and Hall.

(2) Schelkunoff, S. A. Theory of antennas of arbitrary shape and size. *Proc. Inst. Radio Engrs*, *N.Y.*, **30**, 493 (1942).

(3) Smith, R. A. and Holt Smith, C. The elimination of errors from crossed dipole direction finding systems. *J. Instn Elect. Engrs*, **93**, IIIA, 575 (1946).

(4) Essen, L. and Olliver, M. H. Aerial impedance measurements. *Wireless Engr*, **22**, 589 (1945).

Chapter 2

AERIAL ARRAYS

2·0. Introduction

Dipoles, with or without directors or reflectors, are the bricks from which aerial arrays are usually built. In the following chapters we shall be concerned with many different arrangements used to produce aerials to meet various requirements. Fortunately, although theory does not give very reliable values for the self-impedance of a dipole or parasite, it does give reliable values for the mutual impedance between two dipoles. The reason is that whereas the deviation from sinusoidal current in the dipole is a first-order effect as regards self-impedance, it is but a second-order effect as regards the mutual impedance between two dipoles. In calculations of mutual impedance it is therefore in general permissible to assume a sinusoidal distribution of current in both elements.

2·1. Mutual impedance

The mutual impedance between two dipoles is defined as follows. Let e_1 and e_2 be the driving e.m.f.'s in the two dipoles. Then if i_1 and i_2 are the values of the currents at the driving points:

$$e_1 = Z_{11} i_1 + Z_{12} i_2, \quad e_2 = Z_{21} i_1 + Z_{22} i_2. \tag{1}$$

These equations may be written down from the linear nature of the
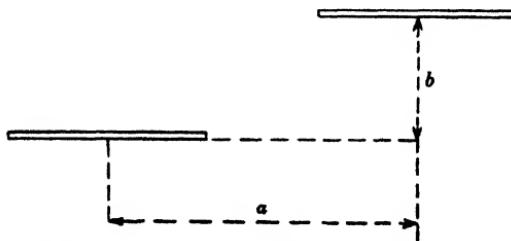
Fig. 6. Relative position of dipoles for evaluation of mutual impedance.

system. Clearly Z_{11} and Z_{22} are the self-impedances of each dipole as may be seen by letting i_1 and $i_2 = 0$ in turn. By a general reciprocity theorem (see § 3·7) $Z_{12} = Z_{21}$ and is known as the mutual impedance between the dipoles. We shall indicate how the quantity Z_{12} is

Table 3. *Mutual impedance $r_{12}+jx_{12}$ of two half-wave dipoles*

b/λ	$a=0$ Resistance r_{12} (ohms)	Reactance x_{12} (ohms)	$a=\frac{1}{2}\lambda$ Resistance r_{12} (ohms)	Reactance x_{12} (ohms)	$a=\lambda$ Resistance r_{12} (ohms)	Reactance x_{12} (ohms)
0·00	73·3	42·2	25·1	19·8	−4·1	−0·7
0·05	71·7	24·3	19·7	10·8	−4·1	−0·6
0·10	67·3	7·6	17·1	3·1	−4·1	−0·4
0·15	60·4	−7·1	13·7	−3·6	−4·1	−0·0
0·20	51·4	−19·1	9·7	−8·6	−4·0	0·5
0·25	40·8	−28·3	4·8	−12·0	−3·9	1·1
0·30	29·2	−34·6	−0·7	−13·1	−3·5	1·7
0·35	17·5	−37·4	−5·7	−13·4	−3·1	2·4
0·40	6·0	−37·6	−10·5	−11·7	−2·4	3·1
0·45	−3·4	−34·8	−13·9	−8·7	−1·6	3·7
0·50	−12·5	−29·9	−16·8	−4·8	−0·7	4·1
0·55	−19·0	−23·4	−18·1	−0·2	0·4	4·2
0·60	−23·2	−16·5	−17·9	4·2	1·5	4·2
0·65	−25·2	−8·0	−16·5	9·1	2·6	3·8
0·70	−24·6	−0·1	−13·7	13·1	3·6	3·1
0·75	−22·5	6·6	−10·0	16·1	4·4	2·3
0·80	−18·5	12·2	−5·6	18·0	5·1	1·1
0·85	−13·0	16·3	−0·8	18·7	5·3	−0·2
0·90	−7·5	18·5	4·0	18·0	5·3	−1·6
0·95	−1·5	19·0	8·5	16·1	4·7	−3·0
1·0	4·0	17·3	12·0	13·1	4·0	−4·2
1·1	12·3	11·2	16·2	5·0	1·5	−5·9
1·2	15·2	2·0	15·3	−3·9	−1·3	−6·0
1·3	12·6	−6·7	10·7	−10·8	−4·5	−4·5
1·4	6·0	−11·8	3·2	−14·0	−6·3	−1·6
1·5	−1·9	−12·6	−4·7	−12·8	−6·0	1·9
1·6	−8·1	−8·4	−10·0	−7·8	−4·3	4·8
1·7	−10·9	−2·0	−11·9	−1·1	−1·2	6·3
1·8	−9·1	4·5	−9·6	5·3	2·5	6·0
1·9	−5·4	8·6	−5·3	9·3	5·1	3·8
2·0	1·1	9·2	0·9	9·7	6·2	0·4
2·1	6·0	6·7	5·9	7·3	5·4	−3·0
2·2	8·2	1·8	8·5	2·6	2·8	−5·3
2·3	7·5	−3·3	7·9	−2·3	−0·5	−5·9
2·4	4·0	−6·8	4·9	−6·2	−3·6	−4·6
2·5	0·9	−7·3	0·4	−6·7	−5·4	−1·8
2·6	−4·8	−5·2	−3·6	−5·7	−5·4	1·6
2·7	−6·7	−1·7	−5·8	−2·8	−3·6	4·1
2·8	−6·3	2·6	−5·5	1·2	−0·7	5·4
2·9	−3·4	5·6	−4·2	4·3	2·4	4·7
3·0	−0·3	5·8	−1·6	5·2	4·2	2·3

(For a and b see fig. 6.)

calculated in Chapter 4. Tables of values are given (tables 3, 4) for two half-wave dipoles placed as shown in fig. 6. The values given in tables 3 and 4 will not vary appreciably with small changes in the length of each aerial when $a > \frac{1}{2}\lambda$ or $b > 0·1\ \lambda$. They are much less sensitive to such changes than the self-impedance.

For full-wave dipoles it is more convenient to use the mutual admittance Y_{12}. Some values are given in table 5.

Table 4. *Mutual impedance $r_{12}+jx_{12}$ of two half-wave dipoles, $b=0$*

a/λ	Resistance r_{12} (ohms)	Reactance x_{12} (ohms)	a/λ	Resistance r_{12} (ohms)	Reactance x_{12} (ohms)
0·50	25·1	19·8	0·80	−0·6	−7·0
0·55	20·3	2·5	0·85	−2·5	−5·6
0·60	14·9	−3·8	0·90	−3·6	−3·9
0·65	9·7	−7·1	0·95	−4·1	−2·2
0·70	5·5	−8·2	1·00	−4·1	−0·7
0·75	2·0	−8·0			

(For a and b see fig. 6.)

Table 5. *Mutual admittance of two full-wave dipoles, $a=0$*

$Y_{12}Z_0^2 = A_{12}+jB_{12}$ ohms; $Z_0 = 120 \log_e [(2l/a)-1]$ ohms.*

| b/λ	A_{12}	B_{12}	b/λ	A_{12}	B_{12}
0·0	197	126	1·1	45	27
0·1	180	21	1·2	51	−6
0·2	132	−54	1·3	39	−33
0·3	71	−96	1·4	12	−51
0·4	3	−102	1·5	−15	−42
0·5	−48	−75	1·6	−36	−24
0·6	−75	−30	1·7	−42	0
0·7	−72	+15	1·8	−30	+21
0·8	−45	45	1·9	−12	36
0·9	−12	63	2·0	+9	33
1·0	+27	51			
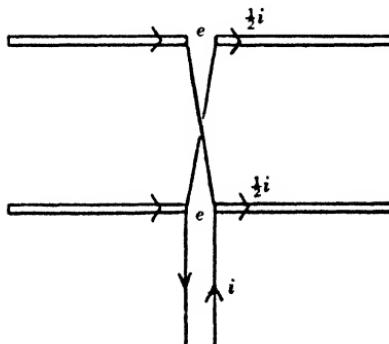
Fig. 7. Two-tier array of half-wave dipoles.

2·2. Impedance of an aerial array

Using the values of self-impedance of the various elements of an array and the mutual impedances between them, the impedance

* a here is clearly the radius of the aerial element.

of the array may be calculated. For example, consider an array made up of two equal half-wave elements fed in phase, as shown in fig. 7, with equal voltages e. The equations (1) become

$$e = Z_{11}\dot{i}_1 + Z_{12}\dot{i}_2, \quad e = Z_{11}\dot{i}_2 + Z_{12}\dot{i}_1.$$

Clearly for symmetry $i_1 = i_2 = \frac{1}{2}i$, where i is the total current entering the aerial. The impedance Z of the aerial is then given by

$$Z = e/i = \frac{1}{2}[Z_{11} + Z_{12}]. \tag{2}$$

2·3. Power-gain of an aerial array

By dividing the power fed into an aerial among a number of elements a greater field strength may be obtained in a particular direction than with a single element. This is simply due to the fact that owing to interference of the radiation from the various elements radiation may be greatly reduced in some directions and in consequence increased in others. The effect of beaming of the radiation by a multi-element array is an example of this effect; a large part of the radiation is concentrated into a relatively small solid angle with consequent gain over an isotropic source. The enhancement of field strength in a given direction, usually the axis of symmetry of the aerial system, is measured by what is known as the 'power-gain' of the aerial. If, for a given input power P, an aerial produces at a point on its axis a certain field strength then, if a power GP were required to produce this field strength with an isotropic source, G is called the 'power-gain' of the aerial. Power-gains are sometimes referred to a half-wave dipole and not to an isotropic source. If this is so it should always be stated. The power-gain relative to an isotropic source will be 1·635 times greater than that relative to a half-wave dipole, as the power-gain of the latter is 1·635 relative to an isotropic source, as we shall see in Chapter 3.

A very elementary consideration gives a rough estimate of the power-gain of an aerial consisting of N elements fed in phase, if we neglect their mutual interactions. Let each element when separately fed with power P produce a field strength E on the axis of the array. Now let the N elements be fed in parallel with power P so that each radiates power P/N. Each will produce at the point in question a field strength $E/\sqrt{N}$. All the N contributions will be in phase and the resultant field will be $\sqrt{N}\,E$, corresponding to what would be obtained from a single element fed with power NP. The power-gain

of the aerial will therefore be N times the gain of each element. This is, of course, only approximately true, as there will in general be mutual interactions between the elements which upset the equi-partition of energy and also change the input impedance of each element.

For example, consider a two-element array fed so as to have equal in-phase currents in each element (fig. 7). Let $Z_{11} = x_0 + jy_0$ and $Z_{12} = x + jy$. Let power P be fed to a single element producing a current of amplitude i at the centre. Let the field strength in a direction normal to the plane of the aerial be ki. Also we have $P = \frac{1}{2}x_0 i^2$. When we have two elements the input impedance of each is clearly $Z_{11} + Z_{12}$, and if i' is the current at the centre we have, since the power P is divided equally between both elements,

$$\tfrac{1}{2}P = \tfrac{1}{2}(x_0 + x)i'^2.$$

The ratio of field strengths given by the two elements and a single element when fed by the same power is $2i'/i$, and the power-gain of the two-element aerial over a single element is

$$4i'^2/i^2 = 2x_0/(x_0 + x). \tag{3}$$

For two elements half a wave apart $x/x_0 = -12 \cdot 5/73$ or $-0 \cdot 17$. The power-gain is therefore $2 \cdot 3$ and not 2 as given by the elementary calculation. The maximum gain is therefore obtained with such a spacing between the elements that x has its greatest negative value, namely, $0 \cdot 65\lambda$. The gain is then nearly $3 \cdot 0$.

By a very general reciprocity relation it may be shown that the power-gain of an aerial for reception is the same as for transmission (see § 3·7). It may then be readily proved that an aerial has the same polar diagram for transmission and reception. If the above theorem holds for a point of observation on the axis of the aerial, by turning the aerial through an angle it may be similarly proved for any other point, if we define the power-gain in a given direction. It follows then that the polar diagrams are the same for transmission as for reception, by comparison with the value in the direction of maximum radiation.

2·4. Polar diagram of aerial arrays

Before considering the polar diagram of multi-element arrays let us recall the form of the polar diagram of the individual elements. The polar diagram of a half-wave dipole is, of course, uniform in its

equatorial plane. In the plane of the dipole it is given approximately by $E = E_0 \cos \theta$, where E_0 is the field strength in the equatorial plane and $\frac{1}{2}\pi - \theta$ is the angle made with the dipole. In fact, this form only holds for an infinitely short or 'Hertzian' dipole, and the polar diagram for a half-wave dipole is given by

$$E = E_0 \cos \left(\tfrac{1}{2}\pi \sin \theta\right)/\cos \theta \qquad (4)$$

(see Chapter 3). This is very closely approximated by the form for a Hertzian dipole, but a little more energy is concentrated towards the equatorial plane. For a full-wave dipole this concentration is more marked and the polar diagram is 'narrower'.
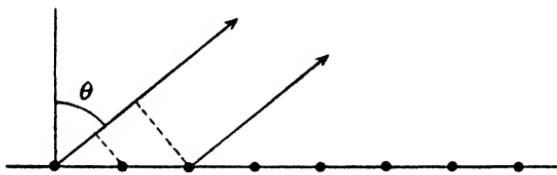
Fig. 8. n-element equi-spaced linear array.

Now suppose we have an aerial made up of n equi-spaced elements in a row as shown in fig. 8. Let each have a polar diagram $f(\theta)$. Let i_r be the magnitude of the current in the rth element and α_r its phase. The field at distance R_r and angle θ from the rth element may then be written

$$E_r = kf(\theta) \exp \left[j(2\pi R_r/\lambda + \alpha_r)\right], \qquad (5)$$

and the field E due to the whole aerial as

$$E = kf(\theta) \sum_{r=1}^{n} i_r \exp \left[j(2\pi R_r/\lambda + \alpha_r)\right], \qquad (6)$$

where we have assumed that the distances R_r are so great that θ is approximately the same for all the elements. We may then write $R_r = R_1 + (r-1)\,d\sin\theta$, where d is the spacing between the elements. When all the α's are zero and the i_r's are equal the series may readily be summed (a simple G.P.). We obtain for the magnitude of E the value

$$|E| = kf(\theta)\frac{\sin\left(n\pi d\sin\theta/\lambda\right)}{\sin\left(\pi d\sin\theta/\lambda\right)} . \qquad (7)$$

This gives the form of the polar diagram of the aerial. The value for $\theta = 0$ is a maximum. Zeroes occur when $\sin\theta = S\lambda/nd$, S being

an integer, with subsidiary maxima in between. If $d < \lambda$ there will only be one principal maximum, but if $d > \lambda$ others will occur when $\sin \theta = S\lambda/d$. When $d < \lambda$ the pattern consists of a main lobe and a series of 'side lobes', as shown in fig. 9. Some examples are also shown in fig. 39. The radiation in the main lobe falls to zero at an angle θ_1 given by

$$\sin \theta_1 = \lambda/nd. \tag{8}$$

When the aperture of the aerial $a = (n-1)d$ is much greater than λ we have the simpler result

$$\theta_1 = \lambda/a. \tag{9}$$

The total width of the beam between zeroes will be $2\lambda/a$ and the width between 3 db. points approximately λ/a, since the variation of field strength over the beam is approximately sinusoidal.

When the currents in the elements are equal the magnitude of the side lobes is appreciable. For example, the main lobe is only $\frac{3}{2}\pi$ or 4·7 times larger than the first side lobe, and the subsequent lobes fall off only as $1/s$. In order to decrease the side lobes, which for some applications can be troublesome, a non-uniform distribution of currents must be used, the current being concentrated towards the centre of the aerial and falling off towards the edges.

One example is the so-called 'biconical array'. If the current i_r in the rth element of an n-element array is proportional to the coefficient of x^r in the expansion of $(1+x)^{n-1}$, the phases α_r being all zero, the sum (6) may again be simply expressed. We have, in fact,

$$|E| = kf(\theta) 2^{n-1} \cos^n \left(\frac{\pi d \sin \theta}{\lambda} \right).$$

This array then produces a polar diagram with a beam whose width decreases as n increases and has no side lobes if $d < \frac{1}{2}\lambda$.

In practice variations from the uniform distribution of current between the elements in order to decrease side lobes generally result in a decrease of gain. However, only in the case of an array of elements spaced $\frac{1}{2}\lambda$ apart is maximum gain obtained with the uniform distribution. It has been frequently assumed in the literature that the maximum gain is obtained with a uniform distribution for any spacing of the elements and even for apertures having a continuous distribution of field strength, but this is not so[1]. It is often said that in order to obtain large gain and a narrow beam

width a large aperture is necessary. In practice this is generally true, but it can easily be shown that it is certainly not true theoretically.

Consider an aerial made up of two parallel half-wave elements spaced a distance d apart and fed in antiphase. Let $d \ll \lambda$. The polar diagram of this array is of the form

$$\text{const.} \times \sin\left(\frac{2\pi d \cos\theta}{\lambda}\right),$$

where θ is the angle with the line of the array. Now construct an end-fire array of two such groups spaced d apart, differing in phase by 180°. This combination will have a polar diagram of the form

$$\text{const.} \times \sin^2\left(\frac{2\pi d \cos\theta}{\lambda}\right).$$

This arrangement is equivalent to an array of three elements with currents proportional to $1, -2, 1$. Now repeat this n times and we have an array with currents proportional to the coefficients of $(1-x)^n$ producing a polar diagram of the form

$$\text{const.} \times \sin^n\left(\frac{2\pi d \cos\theta}{\lambda}\right).$$

Since we have assumed $d \ll \lambda$ this is approximately of the form

$$\text{const.} \times \cos^n\theta.$$

The aperture of the array is nd, and we may always choose nd to be as small as we please even when n is large. In this way we may produce a polar diagram having a beam width as small and gain as large as we please and yet having an aperture as small as we please.*

The aerial described would, of course, not be practicable, as its radiation resistance would be extremely small if d were made small. It would be highly reactive and very difficult to feed. In fact ohmic losses would quite outweigh the energy radiated from the aerial. The above example serves, however, to show clearly that the gain obtainable from a uniformly fed aerial is not the maximum achievable. Improvement on this gain, however, appears to involve the use of current distributions which are impracticable as they lead to an input reactance vastly greater than the radiation resistance[1].

* I am indebted to Mr J. A. Ratcliffe for suggesting this example.

If the phase of each element of the array is advanced by an amount α over the preceding we have $\alpha_r = (r-1)\alpha$ and $|E|$ is given by

$$|E| = kf(\theta) \sin n\left(\frac{\pi d \sin \alpha}{\lambda} - \frac{\alpha}{2}\right) \Big/ \sin\left(\frac{\pi d \sin \theta}{\lambda} - \frac{\alpha}{2}\right). \qquad (10)$$

The principal maximum now occurs at an angle θ_0 given by

$$\theta_0 = \sin^{-1}\left(\frac{\lambda\alpha}{2\pi d}\right). \qquad (11)$$

The beam produced by the aerial has therefore been swung through an angle θ_0.
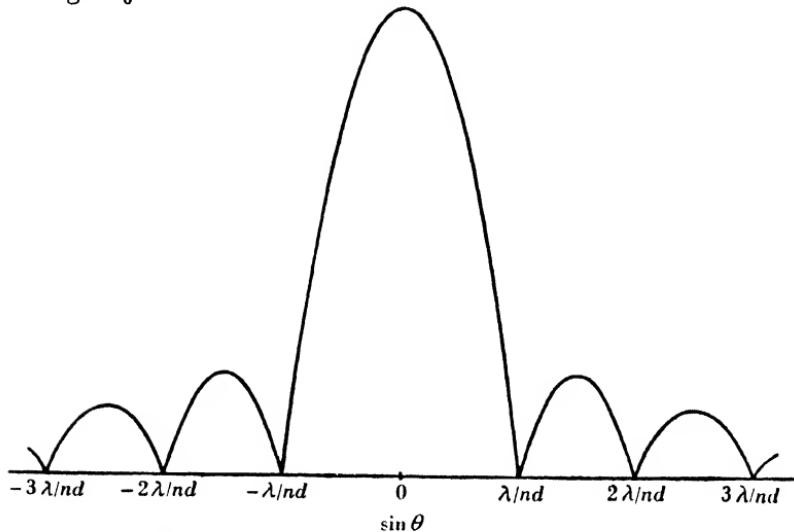
Fig. 9. Interference pattern of multi-element array.

An interesting situation arises when $\alpha = 2\pi d/\lambda$, as then $\theta_0 = \frac{1}{2}\pi$ and the principal maximum occurs along the line of the array. The above condition simply states that the phase of each element is retarded sufficiently to allow its contribution to add in phase to that from all the others along the line of the array. Such arrays are known as 'end-fire' arrays and have been extensively studied[2] (see bibliography in Ref. (2)).

For an aerial consisting of a two-dimensional array of elements a more complex expression for the polar diagram will result. In the principal planes of the aerial, however, two expressions similar to (8) will be obtained, one for each plane.

2·5. Approximate method of estimating the power-gain of an aerial

When we know the beam width in the two principal planes of an aerial we may readily obtain an approximate expression for its power-gain. As stated above the variation of field strength within the beam is approximately sinusoidal in the principal planes. Let α_1, β_1 be the total beam width (to the first zero) in the principal planes. Then if P_0 is the power radiated per unit solid angle the total power in the main beam will be $\alpha_1 \beta_1 P_0$ (since the average of $\sin^2 x$ is $\frac{1}{2}$). If we may neglect the effect of side lobes, compared with an isotropic radiator the aerial will then have power-gain

$$4\pi/\alpha_1\beta_1. \tag{12}$$

If the aerial dimensions in the principal planes a, b are large compared with λ we may use (9) and write (12) in the form

Power gain $= 4\pi ab/\lambda^2$

$= 4\pi \times$ aperture area in square wave-lengths. (13)

This is a general result which may be proved for an aerial which radiates uniformly over its aperture ($\S 3\cdot 5$), but as we have seen may be applied approximately to a large broadside array consisting of uniformly fed elements.

2·6. Effect of earth on aerial polar diagram

Radiation from an elevated transmitter arriving at an aerial near the surface of the earth comes by two paths, first direct and then after reflexion from the earth. The polar diagram in a vertical plane will therefore differ from that for free space since the radiations arriving by the two paths will interfere with each other. Let us consider first a single element whose polar diagram is isotropic in the vertical plane. The direct and reflected rays will have different path lengths as will be seen from fig. 10. Let h be the height of the aerial and ϕ the angle of elevation to a distant source of radiation. From fig. 10 it will be seen that the extra path length Δ will be given by

$$\Delta = 2h\sin\phi.$$

Let the reflected ray have its amplitude reduced by a factor ρ and

its phase delayed by an amount δ on reflexion, then the field strength E due to the combined rays is

$$E = E_0\left[1 + \rho \exp\left\{ j\left(\frac{2\pi\Delta}{\lambda} + \delta\right)\right\}\right], \tag{14}$$

where E_0 is the free-space value. The magnitude of E is given by

$$|E| = E_0\left\{ 1 + \rho^2 + 2\rho \cos\left(\frac{2\pi\Delta}{\lambda} + \delta\right)\right\}^{\frac{1}{2}}. \tag{15}$$

For wave-lengths of the order of 10 m. and horizontal polarization ρ is very nearly equal to 1 and δ to π, so that we have

$$|E| = 2E_0 \sin\left(\frac{2\pi h \sin\phi}{\lambda}\right) \tag{16}$$

on substituting for Δ.
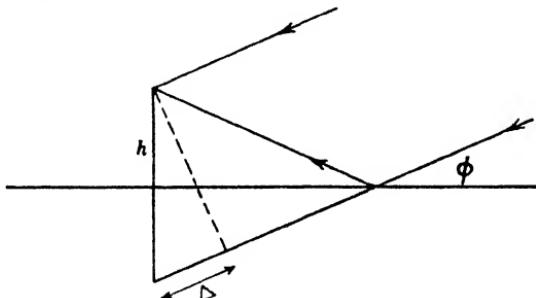
Fig. 10. Direct ray and ray reflected from ground.

The vertical polar diagram will then be given by (16). If the free-space polar diagram has the form $f(\phi)$ the analysis will be modified, and we shall have (taking $\rho = 1$, $\delta = \pi$)

$$E = E_0[f(\phi) - f(-\phi)e^{2\pi j\Delta/\lambda}]. \tag{17}$$

If $f(\phi)$ is symmetrical about the horizontal we shall have $f(\phi) = f(-\phi)$ and

$$|E| = 2E_0 f(\phi) \sin\left(\frac{2\pi h \sin\phi}{\lambda}\right), \tag{18}$$

giving the complete polar diagram. Otherwise we shall have

$$|E| = \left\{ f^2(\phi) + f^2(-\phi) + 2f(-\phi)f(\phi)\cos\frac{4\pi h\sin\phi}{\lambda}\right\}. \tag{19}$$

The form (18) indicates that the vertical polar diagram will have its envelope given by $f(\phi)$, but will consist of a series of maxima

and zeroes, the first maximum occurring at angle ϕ given approximately by

$$\sin \phi_{m1} = \lambda/4h. \tag{20}$$

The first zero ϕ_{01} is given by

$$\phi_{01} = \lambda/2h. \tag{21}$$

The form of the polar diagram is shown in fig. 11. When $f(\phi)$ is not symmetrical a series of maxima and minima will result. In the limit, if $f(\phi)$ is negligible for $\phi < 0$, the free-space polar diagram will be retained.
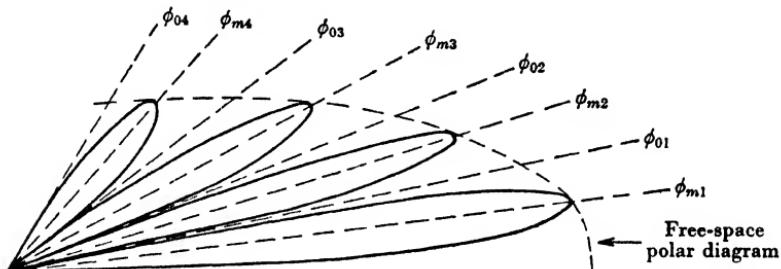
Fig. 11. Vertical polar diagram of aerial above ground.

In the case of vertical polarization the phenomenon associated with the Brewster angle, well known in optics, takes place. At this angle the reflected ray is suppressed. For short waves, maxima and deep minima will be produced (according to height of aerial) just as for horizontal polarization when the angle of elevation is less than the Brewster angle (about $17°$ for land and $6\frac{1}{2}°$ for sea). For angles somewhat greater than the Brewster angle the phase change on reflexion is zero instead of π and the maxima and minima change place.

The above expressions only apply when the curvature of the earth is unimportant. They may, however, be corrected for earth curvature to some extent, but for very small angles of elevation correct diffraction formulae must be used.

2·7. Effect of earth on aerial impedance

The presence of the earth will also affect the impedance of the aerial. To calculate the magnitude of its effect we may replace the earth by an aerial which is the mirror image in the earth of the

original aerial, fed with current just $180°$ out of phase (see, however, §2·6). Using the linear equations relating the currents and voltages in the various elements and taking into account the mutual impedances between the aerial and its image we may calculate the impedance of the aerial in the presence of the earth. For example, consider a horizontal half-wave dipole at height h above a perfectly conducting earth. The e.m.f. E_1 at the terminals will be obtained by putting $i_1 = -i_2 = i$ in (1). The input impedance Z will be then given by

$$Z = Z_{11} - Z_{12}(0, 2h),$$

where $Z_{12}(a, b)$ is the mutual impedance, a, b being defined in fig. 6. The radiation resistance will be $r_{11} - r_{12}(0, 2h)$. It will pass through a series of maxima and minima as h is varied and will tend to zero as h tends to zero, since $\lim_{h \to 0} [Z_{12}(0, 2h)] = Z_{11}$. Numerical values for the impedance of the aerial for any value of h may be obtained from the value taken for Z_{11} and the appropriate value of Z_{12} obtained from table 3. The variation of radiation resistance with h is the same as that shown in fig. 37 (taking $h = d$).

REFERENCES

(1) Woodward, P. M. and Lawson, J. D. On the theoretical precision with which an arbitrary radiation pattern may be obtained from a source of finite size. *J. Instn Elect. Engrs*, **95**, 363 (1948).
(2) Schelkunoff, S. A. Mathematical theory of linear arrays. *Bell Syst. Tech. J.* **22**, 80 (1943).

Chapter 3

ELEMENTARY THEORY OF AERIALS

3·0. Introduction

Most of the properties of simple aerials required by the engineer can be derived by fairly elementary considerations provided some assumption is made regarding the current distribution in the aerial. The calculation of this distribution for all but infinitely thin aerials from the fundamentals of electromagnetic theory is difficult and will be dealt with in the next chapter. For most practical purposes a simplifying assumption is made—that the current distribution in a linear aerial is sinusoidal. This assumption turns out to be a very good approximation for all purposes except for the calculation of the self-impedance of a dipole aerial which, of course, depends very markedly on the current distribution. The approximation also ceases to be valid for aerials many wave-lengths long, when the effect of radiation on the current distribution must be taken into account.

3·1. The sinusoidal theory

In the following treatment we shall therefore assume that the current I at a point distant x from the centre of the aerial is given by an expression of the form

$$I = A e^{(2\pi j x)/\lambda + j\omega t} + B e^{-(2\pi j x)/\lambda + j\omega t}. \tag{1}$$

The constants A and B will be determined by boundary conditions. The form (1) may be compared with similar formulae for the propagation of waves along transmission lines, and the constant λ is interpreted as the wave-length in free space, an approximation only as it turns out, but a good one. For example, if we have an aerial of length $2l$ with ends at $x = \pm l$ fed at the middle the current I will be given by

$$
\begin{aligned}
I &= I_0 \operatorname{cosec} \frac{2\pi l}{\lambda} \sin\left[\frac{2\pi}{\lambda}(l-x)\right] \quad (x>0) \\
&= I_0 \operatorname{cosec} \frac{2\pi l}{\lambda} \sin\left[\frac{2\pi}{\lambda}(l+x)\right] \quad (x<0),
\end{aligned}
\right\} \tag{2}
$$

omitting the time factor $e^{j\omega t}$. Also on the analogy of transmission-line theory the aerial may be said to have what corresponds to the

characteristic impedance. Since, however, except in the case of a biconical aerial, this is not accurately a constant but varies along the aerial, perhaps another name, say 'Wave impedance', would be more appropriate. In the sinusoidal approximation, however, the characteristic impedance, denoted by Z_0, is assumed constant. Z_0 can only be calculated from fairly advanced considerations, as we shall see in the next chapter. It must be emphasized again that the above assumptions cannot be justified *a priori*. Their first justification is that they lead to a simplified treatment which from an engineering point of view has for a long time given results in accordance with practice, but, in fact, their validity under defined conditions is borne out by fundamental theory. In the course of the present chapter we shall derive a few results which do *not* depend on the sinusoidal assumption. We shall make it clear when the result depends on this assumption and is therefore restricted in its validity.

3·2. Field strength due to an aerial of power-gain G

We shall consider the field strength at a point distant R from the aerial in the direction of maximum radiation. We shall take R to be sufficiently great so that we are concerned only with the radiation field, which falls off as $1/R$. For an isotropic source, radiating power P, the power density at distance R will be $P/4\pi R^2$. For an aerial of power-gain G it will be $GP/4\pi R^2$. Now if E is the field strength (amplitude) the power density is given by $E^2/240\pi$ (m.k.s. units will be used throughout) (see Ref. (1)). We have therefore

$$\frac{E^2}{240\pi} = \frac{GP}{4\pi R^2}$$

or
$$E = (60)^{\frac{1}{2}} G^{\frac{1}{2}} P^{\frac{1}{2}}/R. \tag{3}$$

E is expressed in volts/metre, P in watts, and R in metres.

3·3. Polar diagram and power-gain

The expression (3) gives the field strength in the direction of maximum radiation. The field strength in any other direction is obtained from the polar diagram. When this is known the power-gain G may, of course, be calculated at once. Let $f(\theta, \phi)$ represent the polar diagram. Then the power density at distance R for any direction (θ, ϕ) is given by

$$W(\theta, \phi) = W_0 f^2(\theta, \phi), \tag{4}$$

where W_0 is the power density in the direction of maximum radiation. By definition of power-gain we have, equating the total flux of power to $1/G$ times that from an isotropic source radiating with power density equal to that at the maximum,

$$4\pi W_0/G = \iint W_0 f^2(\theta, \phi)\, d\omega, \quad 4\pi G^{-1} = \iint f^2(\theta, \phi)\, d\omega. \tag{5}$$

This is an important general result which we have already applied in a special form in § 2·5.

3·3·1. *The Hertzian dipole*

The Hertzian dipole is an elementary electric doublet of dimensions small compared with a wave-length. All elements of the current in the dipole are considered as in phase and the field strength

(a) Hertzian (electric) doublet   (b) Magnetic doublet
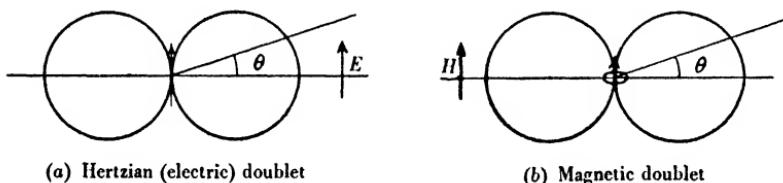
Fig. 12. Polar diagrams of elementary doublets.

is then proportional to the sine of the angle which the direction in question makes with the doublet. The polar diagram is then given by

$$f(\theta) = \cos\theta, \tag{6}$$

where θ is the angle between the direction of observation and a line at right angles to the doublet (fig. 12 a).

The power-gain is given by

$$G^{-1} = \frac{1}{4\pi} \int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi} \int_0^{2\pi} \cos^3\theta\, d\theta\, d\phi$$

or
$$\underline{G = \tfrac{3}{2}}. \tag{7}$$

Similarly, of course, we can have an elementary magnetic doublet consisting of a small-plane coil of dimensions much smaller than a wave-length having a similar polar diagram but with polarization turned·through a right angle, the axis of the doublet being normal to the plane of the coil (fig. 12 b).

3·3·2. *Half-wave dipole*

The polar diagram of a half-wave dipole, as we have seen, does not differ much from the form (6). However, when we take into account the phase delays of the radiations from successive elements along the dipole, a slightly different form is found. We must first write down an expression for the field radiated by a current element. When the distance R is sufficiently great, so that we are only concerned with the radiation field, the electric field ΔE is transverse and is given by (2),

$$\Delta E = \frac{60\pi \cos\theta}{\lambda R} \Delta x \, I \exp\left\{-j\left(\frac{2\pi R}{\lambda} - \omega t - \frac{\pi}{2}\right)\right\}, \tag{8}$$

the current having length Δx and carrying a current $Ie^{j\omega t}$. Omitting the time factor and taking into account the phase delays from the
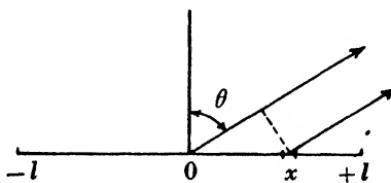

Fig. 13. Radiation from linear dipole.

various elements of the dipole the field radiated from a dipole carrying a current given by (2) will be (see fig. 13)

$$E = \frac{60\pi \cos\theta}{\lambda R} \int_{-l}^{l} I(x) \, e^{2\pi j x \sin\theta/\lambda} \, dx, \tag{9}$$

where $I(x)$ is given by (2).

When $l = \frac{1}{4}\lambda$ we have a half-wave aerial and (9) simplifies to

$$E = \frac{60\pi I_0 \cos\theta}{\lambda R} \int_{-\frac{1}{4}\lambda}^{\frac{1}{4}\lambda} \cos\frac{2\pi x}{\lambda} \, e^{(2\pi j x \sin\theta)/\lambda} \, dx,$$

giving $$|E| = \frac{60 I_0}{R} \frac{\cos\left(\frac{1}{2}\pi \sin\theta\right)}{\cos\theta}. \tag{10}$$

If the expression (10) is plotted out as a function of θ it will be seen not to differ much from the approximate form

$$E = \frac{60 I_0}{R} \cos\theta, \tag{11}$$

which, as we have seen, is frequently sufficiently accurate. Comparing (11) with (8) we see that the effective length of a half-wave

dipole may be taken as λ/π, i.e. the dipole is equivalent to an elementary doublet of this length carrying a current I_0. The reciprocal property of an aerial for transmitting and receiving (which we shall prove in § 3·7) may be used to derive from this result an expression for the e.m.f. induced in a half-wave dipole. An elementary doublet of length Δx clearly has an e.m.f. $E_0 \Delta x$ induced in it by a field E. A half-wave dipole therefore has an e.m.f. of amplitude E_0 given by

$$E_0 = E\lambda/\pi. \tag{12}$$

The power-gain of a half-wave dipole may now be calculated. It is given by

$$G^{-1} = \frac{1}{4\pi} \int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi} \int_0^{2\pi} \frac{\cos^2(\frac{1}{2}\pi \sin\theta)}{\cos^2\theta} \cos\theta \, d\theta \, d\phi. \tag{13}$$

This integral may be readily evaluated numerically and gives

$$G = 1 \cdot 635 \; .$$
$$= \tfrac{3}{2} \times 1 \cdot 09. \tag{14}$$

The gain of a half-wave dipole over an isotropic source is therefore $1\cdot635$, and over a Hertzian dipole is $1\cdot09$. All the results given in § 3·3·2 depend, of course, on the sinusoidal approximation and are only valid to a corresponding accuracy.

3·4. Radiation resistance

Since a large part of the power fed to a resonant dipole is radiated it is clear that its input impedance must have a resistive component related to the loss of power by radiation. This is known as the radiation resistance. In fact, for the efficient dipoles used in short-wave work, ohmic losses in the dipole and in the surrounding media such as the earth are comparatively small, and we may calculate the resistive part of the input impedance from the radiation alone. Let r_0 be the resistive part of the input impedance. In the case of a resonant dipole r_0 will be equal to the whole input impedance. Then we have, if i_0 is the amplitude of the current at the input,

$$\text{Power dissipated} = \tfrac{1}{2} r_0 i_0^2. \tag{15}$$

We can also equate this to the flux of energy through a large sphere of radius R. We have then

$$\tfrac{1}{2} r_0 i_0^2 = \frac{1}{240\pi} \iint E^2(\theta, \phi) R^2 d\omega. \tag{16}$$

If E_m is the value of E in the direction of maximum radiation, we have

$$\tfrac{1}{2}r_0 i_0^2 = \frac{R^2 E_m^2}{240\pi} \int\int f^2(\theta, \phi)\, d\omega^2,$$

where $f(\theta, \phi)$ is the polar diagram function. Using (5) this reduces to

$$\tfrac{1}{2}r_0 i_0^2 = R^2 E_m^2/60G. \tag{17}$$

For a Hertzian dipole

$$R^2 E_m^2 = 3600\pi^2 \Delta x^2 i_0^2/\lambda^2.$$

So we have, since $G = \tfrac{3}{2}$,

$$r_0 = 80\pi^2 \Delta x^2/\lambda^2 \text{ ohms}. \tag{18}$$

For a half-wave dipole the effective value of Δx is, as we have seen, λ/π, so that $r_0 = 80$ ohms to the approximation that the polar diagram is given by $\cos\theta$. A much better approximation is, however, obtained by taking $1\cdot635$ for the power-gain as derived from the polar diagram given by (10). We have then from (17), using (10) for E_m with $\theta = 0$,

$$r_0 = 120/G \text{ ohms}. \tag{19}$$

Using the value $G = 1\cdot635$ we have $r_0 = 73\cdot2$ ohms. This is the value usually quoted. It is correctly deduced from the assumption of a sinusoidal current. It may be shown (see Chapter 4) that as an aerial is made thinner this assumption becomes more and more nearly correct, so that the value $73\cdot2$ ohms is the limit to which the radiation resistance tends as the diameter of the aerial wire tends to zero.

Frequently it is desired to feed a short-wave aerial with an unbalanced feeder, for example a concentric feeder, and in this case a 'unipole' or half-dipole is used, one end being near a metal sheet which forms an 'earth'. The image of the 'unipole' in the sheet, which must be a very good conductor, forms the other half of the dipole. The first resonant length of the unipole will then be approximately $\tfrac{1}{4}\lambda$. For long waves this arrangement is generally used, the earth itself acting as the metal sheet. For short waves, however, the earth is by no means a perfect conductor, and the image instead of being in phase is more nearly in anti-phase, as we have seen, and the following arguments do not apply unless an artificial earth, made of conducting wires, is used. The polar diagram of the unipole and its image will be the same as for a dipole, but all the

energy fed to the aerial will be radiated into a hemisphere. From (16), therefore, it will be seen that the radiation resistance will have half the value given by (18) and (19). For example, for an aerial of effective height h above a perfectly conducting earth, the radiation resistance will be (writing $\Delta x = 2h$) $160\pi^2 h^2/\lambda^2$ ohms. For a resonant $\frac{1}{4}\lambda$ unipole the radiation resistance will be about 37 ohms.

It will be appreciated that only the resistive part of the input impedance has been calculated. Unless the aerial is of length such that it behaves as a resonant $\frac{1}{2}\lambda$ dipole (or $\frac{1}{4}\lambda$ unipole) the reactive part will have to be obtained by other means. If we try to calculate the radiation resistance of a full-wave dipole by the above argument we run into trouble, since according to the sinusoidal approximation the input current to a full-wave dipole is zero. In fact, the limit to which the anti-resonant resistance tends as the dipole becomes infinitely thin (and the sinusoidal approximation becomes exact) is infinity. It is not therefore surprising that the above method of calculation fails in this case, and a more elaborate treatment based on more fundamental considerations is required. That the impedance is high, becoming greater as the aerial is made thinner, may, however, be expected.

3·5. Power absorbed by a receiving aerial

Suppose we have a receiving aerial whose input impedance is $r_0 + jx_0$. Let the amplitude of the e.m.f. induced in the aerial by a field $Ee^{j\omega t}$ parallel to it be e_0.

Now suppose the aerial feeds into a load $r_1 + jx_1$ as shown in fig. 14. The power dissipated in the load will then be

$$\frac{1}{2}\frac{r_1 e_0^2}{(r_1 + r_0)^2 + (x_1 + x_0)^2},$$

and in the aerial (re-radiated)

$$\frac{1}{2}\frac{r_0 e_0^2}{(r_1 + r_0)^2 + (x_1 + x_0)^2}.$$

If the aerial is matched to its load $r_1 = r_0$ and $x_1 = -x_0$, and we have for P_a the power absorbed (equal to the power re-radiated)[*]

$$P_a = \frac{1}{8}\frac{e_0^2}{r_0}. \tag{20}$$

[*] This only gives the power re-radiated if the current distribution for receiving is the same as for transmitting. This, in general, is not so—cf. p. 29.

In the case of a Hertzian dipole of length Δx, $e_0 = E\Delta x$, and substituting for r_0 from (18) we have

$$P_a = E^2\lambda^2/640\pi^2. \tag{21}$$

Now if we suppose the aerial to have an 'effective absorbing area' A_H, P_a will be equal to $E^2 A_H/240\pi$ (area times power flux). Equating these we have, for the absorbing area of a Hertzian dipole,

$$A_H = 3\lambda^2/8\pi. \tag{22}$$

It is interesting to note that A_H is independent of the length Δx of the dipole.
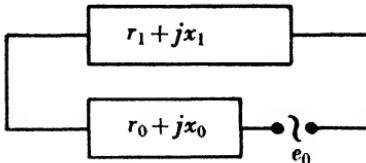

Fig. 14. Equivalent circuit of receiving aerial.

For an aerial having gain G relative to an isotropic source and therefore $\frac{2}{3}G$ relative to a Hertzian dipole we have clearly the absorbing area equal to $\frac{2}{3}G$ times A_H or

$$A = \lambda^2 G/4\pi. \tag{23}$$

This is a very important fundamental equation. Note that it does not depend on any approximation.

The gain, when transmitting, of an aerial having an aperture of area A, large in both dimensions compared with λ, and uniformly illuminated is given by*

$$G = \frac{4\pi A}{\lambda^2}. \tag{24}$$

The 'absorbing area' is therefore just equal to A, so that *all* the radiation falling on the aperture is fed to a matched load.

---

* This formula may also be deduced from Huygens's Principle (see Ref. 5, Ch. V). The field at large distance R from the aperture at a point P on the axis is

$$E = \frac{1}{\lambda} \int\!\!\int_A \frac{E_a e^{2\pi i x/\lambda}}{x}\, ds,$$

x being the distance from the element ds to P. If E_a is constant over the aperture, when R is large this reduces to

$$\frac{AE_a}{\lambda R} e^{2\pi i R/\lambda}.$$

The radiated power is $E_a^2 A/240\pi$ and if radiated uniformly would produce a field at distance R of amplitude equal to $A^{\frac{1}{2}}E_a/(4\pi)^{\frac{1}{2}} R$. Comparing these expressions we obtain for the gain

$$G = 4\pi A/\lambda^2.$$

For a half-wave dipole the effective aperture $A_{\lambda/2}$ is found from (23) to be given by

$$A_{\lambda/2} = \frac{\lambda^2 . 1 \cdot 635}{4\pi} \simeq \tfrac{1}{8}\lambda^2.   (25)$$

It is interesting to note that if the aerial is a parasite, and so is continuous at the centre, and resonant, we may take $r_1 = x_1 = x_0 = 0$. The power absorbed by the aerial and re-radiated is then

$$\tfrac{1}{2}e_0^2/r_0,   (26)$$

i.e. four times that for a matched aerial. If it is not resonant the amount will be

$$\tfrac{1}{2}r_0 e_0^2/(r_0^2 + x_0^2).   (27)$$

3·6. Radiation from aerial apertures

So far we have regarded the radiation from an aerial as coming from a number of elements carrying oscillating currents. For some purposes, however, it is more convenient to regard the radiation in terms of Huygens's Principle as arising from a distribution of field strength across an aperture. Let us use this conception to deduce a general expression for the polar diagram of the radiation from such an aperture. Let us first consider a rectangular aperture for simplicity and work out the polar pattern in one of the principal planes. According to Huygens's Principle* the field strength at a distant point will be the sum of the contributions of all the elements of the aperture, each contributing an amount proportional to the field strength there. The constant of proportionality need not concern us here. Let x be the distance of an element from one of the sides of the aperture as shown in fig. 15, and θ the angle made by the direction in which the radiation is being observed, with the normal to the aperture. Let $E_a(x)$ be the electric field in the plane of the aperture, polarized, we shall suppose, at right angles to the plane in which we wish to calculate the polar pattern. Then we have by Huygens's Principle the distant field E given by

$$E(\theta) = k \int_0^a E_a(x) e^{2\pi j x \sin\theta/\lambda} dx,$$

where a is the width of the aperture.

Now if we consider $\sin\theta$ as the variable instead of θ, and measure x in wave-lengths, we note that $E(\theta)$ is just the Fourier transform

* Ref. 5, Ch. V.

of the aperture distribution of field strength, and the polar diagram is just the spectrum in terms of $\sin\theta$ of this distribution. From the well-known properties of Fourier spectra of pulses several important results follow at once (5). A uniform distribution corresponds to a square-topped pulse, and its polar diagram is given by

$$E(\theta) = \text{const.} \frac{\sin(\pi a \sin\theta/\lambda)}{\pi a \sin\theta/\lambda}. \qquad (28)$$

This is very similar to the polar pattern obtained for a linear array and given by equation (7) of Chapter 2. Its form is also given by fig. 9 with appropriate scale. The first zero is given by $\theta = \lambda/a$ as before. Just as the spectrum of a square-topped pulse contains high harmonics due to the discontinuities at the beginning and end,
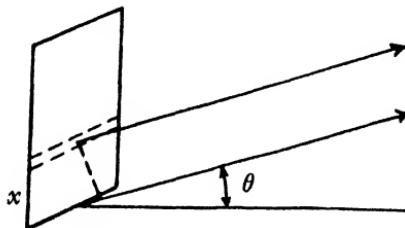

Fig. 15. Field radiated from aerial aperture.

so the uniform aperture distribution gives large side lobes. By rounding off a pulse the high harmonics are reduced, and by tapering off an aperture distribution towards the edges the side lobes are reduced.

The above analysis is of the utmost importance for aerials which obtain a large aperture by using parabolic reflectors or horns, such as are widely used for the microwaves. We shall not be much concerned with such aerials in this book,* and have only included this generalization of the analysis given in Chapter 2 in order to show the generality of the principles involved.

3·7. The reciprocal property of transmitting and receiving aerials

Let us recall first a well-known theorem connected with four-terminal linear networks. If the input voltage and current are e_1, i_1

* They will be dealt with in another of the present series.

at the pair of terminals marked (1) (fig. 16) and e_2, i_2 are the corresponding quantities at the pair of terminals marked (2), then

$$e_1 = Z_{11}i_1 + Z_{12}i_2, \quad e_2 = Z_{21}i_1 + Z_{22}i_2, \tag{29}$$

simply by the *linear* property of the system, the quantities Z_{11}, etc., being independent of the currents and voltages. The reciprocity theorem for four-terminal networks consists of proving[3] that in fact there are not four constants $Z_{11}, Z_{12}, Z_{21}, Z_{22}$ describing a general four-terminal network but only three, since $Z_{12} = Z_{21}$. It follows at once by putting in turn $i_1 = i_0$, $i_2 = 0$ and $i_2 = i_0 = 0$, that if a current i at (1) produces an open-circuit e.m.f. e at (2) then a current i at (2) produces the same open-circuit e.m.f. e at (1).
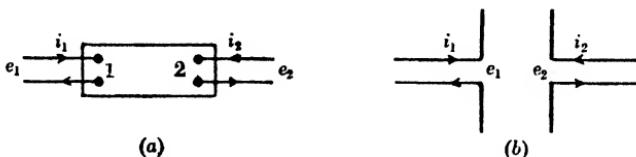

Fig. 16. Transmitting and receiving aerials as four-terminal network.

The properties of a transmitting and receiving antenna system may be described in terms of a four-terminal network provided no non-linear elements enter into the system. The properties of the network may be entirely described in terms of the quantities Z_{11}, Z_{22}, known as the self-impedances, and Z_{12} the transfer impedance or mutual impedance. When we know these we need to know nothing of what happens between the terminals except that all processes are linear. Now in the case of long waves there is some doubt that the transmission between two aerials is invariably a linear process. Ferro-magnetic materials in the earth behave in a non-linear fashion as does the ionosphere under certain circumstances. For very short waves, however, there is little doubt that the transmission process can be taken as linear.

Certain important properties of aerials follow at once. The impedance Z_t of an aerial for transmission is given by the e.m.f. across the input terminals divided by the input current. Since all receiving aerials are assumed to be very distant we may take Z_t equal to Z_{11}. If we now use the aerial (2) for reception the open-circuit e.m.f. e induced in the aerial by a current i_1 in a distant

transmitting aerial (1) will be $e = Z_{12}i_1$. If we now connect a load Z across its terminals we have a current i_2 given by the equation

$$e_2 = Z_{22}i_2 + Z_{12}i_1.$$

But $e_2 = -Zi_2$ and $Z_{22} = Z_r$, the self-impedance of the receiving aerial. We therefore have

$$i_2 = -e/(Z_r + Z). \tag{30}$$

(The negative sign simply corresponds to the difference in direction of current for transmission and reception.)

This result could, of course, have been obtained at once from Thévenin's theorem, regarding the receiving aerial as a generator of e.m.f. e and impedance Z_r. What we have wished to emphasize is the equivalence of Z_t, Z_r for transmitting and receiving. If we reversed the aerials, Z_r, being equal to Z_{22}, would be used also for the impedance when transmitting. If we know e and Z, Z_r could of course be determined from (30), but in general it is determined, for the transmitting condition, for example, by connecting a transmission line to the aerial and observing the standing wave pattern when the aerial is excited through the transmission line. Although equation (30) is frequently used, the fact that the Z_r in it is the same as that for transmission is not always appreciated, and the dependence of the validity of this fact on the linearity of the system is not always apparent.

We may now derive the result that the power-gain of an aerial for reception is the same as that for transmission. Consider first transmission and reception between two aerials (0) and (1) and compare with transmission and reception over the same path between two aerials (0) and (2). For (0) and (1) we have

$$e_0 = Z_{00}i_0 + Z_{01}i_1, \quad e_1 = Z_{01}i_0 + Z_{11}i_1. \tag{31}$$

For (0) and (2) we have

$$e_0 = Z_{00}i_0 + Z_{02}i_2, \quad e_2 = Z_{02}i_0 + Z_{22}i_2. \tag{32}$$

Now let a current i be fed in turn to aerials (1) and (2) and the resulting signal be received by aerial (0) which is open-circuited. Let e_{01} and e_{02} be the open-circuit e.m.f.'s induced. Then we have, putting $i_0 = 0$, $i_1 = i_2 = i$ in (31) and (32),

$$e_{01} = Z_{01}i, \quad e_{02} = Z_{02}i.$$

Now, by definition of power-gains G_{t1}, G_{t2} of aerials 1 and 2 for transmitting,
$$|e_{01}|^2/|e_{02}|^2 = P_1 G_{t1}/P_2 G_{t2}, \qquad (33)$$

where P_1 and P_2 are the powers fed to (1) and (2). Let R_1 and R_2 be the real parts of Z_{11} and Z_{22}, then $P_1/P_2 = R_1/R_2$. From (33) we then have
$$G_{t1}/G_{t2} = R_2 |Z_{01}|^2/R_1 |Z_{02}|^2. \qquad (34)$$

Now let a current i_0 be fed to aerial (o), aerials (1) and (2) acting as receivers. The open-circuit e.m.f.'s e_1 and e_2 received at the aerials will be given by
$$e_1 = Z_{01} i_0, \quad e_2 = Z_{02} i_0. \qquad (35)$$

Note therefore that $e_{01}/e_{02} = e_1/e_2$. This is one way of expressing the reciprocal property. However, it is not the usual one. We must define with some care what we mean by the power-gain of a receiving aerial. It is the ratio of maximum power that can be fed from the aerial to a load to that which can be fed by an aerial having an isotropic polar diagram, when receiving under the same conditions. Note that it is certainly *not* the ratio of squares of open-circuit e.m.f.'s. (For example, a Hertzian dipole has gain independent of its length but open-circuit e.m.f. proportional to its length.)

The maximum power will be transferred to a load by the receiving aerials (1), (2) when they are matched to the load and will be proportional to $|e_1|^2/R_1$ and to $|e_2|^2/R_2$. We therefore have, for the power-gains for receiving G_{r1}, G_{r2},
$$G_{r1}/G_{r2} = |e_1|^2/R_1 \div |e_2|^2/R_2. \qquad (36)$$

From (35) we therefore have
$$G_{r1}/G_{r2} = R_2 |Z_{01}|^2/R_1 |Z_{02}|^2. \qquad (37)$$

Comparing (34) and (35) and taking one of the aerials as standard, we have the result that the power-gain for reception is the same as for transmission. As explained in §2·3 it follows at once that the polar diagram of an aerial is the same for transmission as for reception.

An interesting application of a reciprocity relation is its use to derive an expression for the open-circuit e.m.f. produced in a linear aerial by any incident field. One form of a general reciprocity relation for any linear electrical system is as follows[3]. Let e.m.f.'s $e_1, e_2, ..., e_r, ...$ be applied at points of the system producing at these

points the currents $i_1, i_2, ..., i_r, ...$. Let another set of e.m.f.'s $e'_1, e'_2, ..., e'_r, ...$ produce currents $i'_1, i'_2, ..., i'_r, ...$. Then we have the relationship

$$\Sigma i_r e'_r = \Sigma i'_r e_r. \tag{38}$$

Now let x be the distance measured along a linear aerial from its feed point. Let any distribution of field $E_1(x)$ produce a current distribution $i_1(x)$, and another $E_2(x)$ a current distribution $i_2(x)$. We may write the e.m.f. applied to an element dx of the aerial as $E_1(x) dx$, etc., and (38) may then be written in the form

$$\int E_1(x) i_2(x) dx = \int E_2(x) i_1(x) dx, \tag{39}$$

where the integration is taken along the aerial.

In particular, if $E_1(x)$ represents the transmitting condition when the aerial is being driven at its centre, $E_1(x) = 0$ except near $x = 0$, and so we have

$$ei_{02} = \int E_2(x) i_t(x) dx, \tag{40}$$

where e is the voltage at the input, equal to $Z_0 i_0$, $i_t(x)$ being the current distribution when transmitting and having the value i_0 at $x = 0$. i_{02} is the value of $i_2(x)$ at $x = 0$. If the aerial is receiving a field given by $E_r(x)$ we then have the interesting relationship

$$i_{0r} = \frac{1}{Z_0 i_0} \int E_r(x) i_t(x) dx, \tag{41}$$

where i_{0r} is the current produced at the input when short-circuited. The open-circuit e.m.f. e_r is, of course, equal to $Z_0 i_{0r}$, and is given by

$$e_r = \frac{1}{i_0} \int E(x) i_t(x) dx. \tag{42}$$

If $E(x) = E$ a constant, we have

$$\frac{e_r}{E} = \frac{1}{i_0} \int i_t(x) dx. \tag{43}$$

If $x = 0$ is at a current maximum we normally call the quantity on the left-hand side the 'effective length of the aerial for receiving' h_r. The e.m.f. produced by the field is just Eh_r. The quantity on the right is what is known as the 'effective length for transmitting', h_t, i.e. the length of a Hertzian dipole radiating the same power in its

equatorial plane when carrying the same current as the aerial at its feed point. Equation (43) shows, as one would expect, that h_l and h_r are equal[4].

3·8. Power received from a distant transmitting aerial

The theory given in the foregoing sections enables us to calculate the power received by an aerial from a distant transmitter. We can, in fact, calculate the transfer impedance Z_{12} defined in (29). Let P be the power fed to the transmitting aerial and G its gain. The field strength E at distance R is then given by (3). The power density absorbed by the receiving aerial if correctly matched is $AE^2/240\pi$, where A is its absorbing area. Now from (23) we have, for P_r the received power,

$$P_r = \frac{G_r \lambda^2}{4\pi} \frac{E^2}{240\pi} = \frac{PG_r G_l \lambda^2}{16\pi^2 R^2}. \tag{44}$$

The current i_t in the transmitting aerial is given by $P = \frac{1}{2} r_{0t} i_t^2$, r_{0t} being the real part of the input impedance of the aerial. The open circuit e.m.f. in the receiving aerial is then $Z_{12} i_t$, and the power fed to a matched load is, from (20), $|Z_{12}|^2 i_t^2 / 8 r_{0r}$, r_{0r} being the real part of the self-impedance of the receiving aerial. Hence

$$\frac{|Z_{12}|^2 i_t^2}{8 r_{0r}} = \frac{1}{2} r_{0t} i_t^2 \frac{G_r G_l \lambda^2}{16\pi^2 R^2};$$

therefore

$$|Z_{12}|^2 = \frac{r_{0r} r_{0t} G_r G_l \lambda^2}{4\pi^2 R^2}. \tag{45}$$

The above calculation applies only to free space. To include the effect of the presence of the earth (44) will have to be multiplied by the square of a factor such as occurs in equation (18) of Chapter 2.

REFERENCES

(1) Stratton, J. A. *Electromagnetic Theory*, p. 16 (1941). McGraw Hill.
(2) Stratton, J. A. *Electromagnetic Theory*, p. 440.
(3) Shea, T. E. *Transmission Networks and Wave Filters* (1929). London: Chapman and Hall.
(4) Burgess, R. E. The relationship between the transmitting and receiving characteristics of aerials. *Wireless Engr*, **21**, 154 (1944).
(5) Smith, R. A. Electronics, Chapter VI (Radar). Pilot Press (1947).

Chapter 4

FUNDAMENTAL THEORY OF RADIATION FROM LINEAR AERIALS

4·0. Introduction

The problem of calculating theoretically the field radiated from a linear aerial driven in a given manner is a peculiarly difficult one. The cylindrical aerial most commonly used is particularly difficult to treat exactly, and the only really satisfactory mathematical treatments given so far are for aerials in the shape of a spheroid or two coaxial cones. If the field produced by the aerial is calculable, so also is the current in the aerial. Much of the early work on aerials (1, 2, 3), however, assumed a particular distribution of current in the aerial, namely, a sinusoidal distribution. Such an assumption, made *a priori*, is, of course, completely unjustified from a theoretical standpoint, but happens to be very accurately true in practice in the case of thin aerials. This assumption led to what is known as the 'sinusoidal theory' of aerials. This theory, though based originally on very shaky foundations, in fact gives surprisingly good results in practice. As we shall see, it may in fact be proved that as the thickness of an aerial tends to zero the current tends to sinusoidal form, and a proper theoretical treatment shows that departures from this form are small so long as $\log_e 2l/a$ is large compared with unity, $2l$ being the length of the aerial and a its radius.

The implications of the sinusoidal theory and its imperfections have been discussed by Schelkunoff(4). The assumption of a sinusoidal distribution of current leads to a non-zero tangential electric field at the surface of the aerial, assumed a perfect conductor, in violent contradiction with the boundary condition which maintains that the tangential component of electric field should be zero there. It is generally assumed that the radiation flowing from an aerial carrying such a sinusoidal current corresponds to what happens in practice, but this is certainly not the case. If we calculate the rate of flow of radiation by means of the Poynting vector, using the field derived from the sinusoidal current, we are led to believe that the energy flows from the surface of the aerial all along its length. This at first sight seems to tie up with the idea of each current element

radiating, but, in fact, is far from the truth. If the aerial is a perfect conductor the tangential field at its surface must be *zero*, so that no energy can flow from its surface except at the point where it is driven. If it is an imperfect conductor the flow must be *into* its surface. The energy, in fact, all flows from the source and flows along the aerial, i.e. is guided by it. Energy then flows out radially from the driving point and is reflected when it comes to the ends of the aerial, giving the well-known result that very little radiation takes place along the length of the aerial. In other directions some of the energy flows off into space, mostly in a direction at right angles to the aerial.

4·1. The 'sinusoidal' theory of aerials

Since the sinusoidal theory of aerials gives results which are of great importance from an engineering point of view, we shall consider it in some detail, returning later to its theoretical justification. As we have seen it begins by assuming that an aerial carries a sinusoidal current in the form

$$i(x) = A \sin \frac{2\pi}{\lambda} (l - x) \quad (x > 0)$$

$$= A \sin \frac{2\pi}{\lambda} (l + x) \quad (x < 0), \tag{1}$$

which is zero at the free ends $x = \pm l$.

The field due to this current is then calculated by means of electromagnetic potentials (Ref. (1) of Chapter 3). Let $i(x)$ be the current and $q(x)$ be the charge per unit length, then the vector and scalar potentials $\mathbf{A}(x)$ and $\phi(x)$ are given by

$$\mathbf{A} = \frac{\mu}{4\pi} \int [i(x)] \, \mathbf{ds}/r, \tag{2}$$

$$\phi = \frac{1}{4\pi\epsilon} \int [q(x)] \, ds/r. \tag{3}$$

The square brackets denote retardation and r is the distance from the point at which $\mathbf{A}$ and ϕ are to be evaluated from the element ds. The field is obtained from $\mathbf{A}$ and ϕ by means of the equations

$$\mathbf{E} = -\operatorname{grad} \phi - \frac{\partial \mathbf{A}}{\partial t}, \tag{4}$$

$$\mathbf{H} = \frac{1}{\mu} \operatorname{curl} \mathbf{A}. \tag{5}$$

For a field varying as $e^{j\omega t}$ **A** is given for a linear current (apart from the time factor) by

$$\mathbf{A} = \frac{\mu}{4\pi} \int \dot{i}(x) \frac{e^{-jkr}}{r}\, dx, \tag{6}$$

where $k = 2\pi/\lambda$.

The crudest form of the sinusoidal theory simply takes over an analogy from the theory of transmission lines. The assumption of a sinusoidal distribution of current leads one to attribute to the line a uniform characteristic impedance Z_0. Proceeding as in transmission-line theory one may define an inductance per unit length by L assuming that L times the temporal rate of change of current at a point gives the electric field resulting from the current. To calculate L retardation is neglected, the contribution being assumed to come from neighbouring points only. Proceeding in this way we shall have from (6)

$$E_1 = -\frac{\partial A}{\partial t} \tag{7}$$

as the contribution to the field from the *current* or

$$E_1(x') = j\omega i(x') \frac{\mu}{4\pi} \int \frac{dx}{r}. \tag{8}$$

The function $i(x)$ is taken outside the integral on the assumption that the important contribution comes from neighbouring points only. Hence

$$L = \frac{\mu}{4\pi} \int \frac{dx}{r}. \tag{9}$$

Using the normal transmission-line relation we may define a 'characteristic impedance' $\zeta_0 = 120\pi L/\mu$, or

$$\zeta_0(x') = 30 \int \frac{dx}{r}. \tag{10}$$

For a cylindrical aerial of radius a, $r^2 = (x - x')^2 + a^2$. When we are dealing with balanced aerials we are not normally concerned with impedances to earth but between two parts of the aerial, and in the literature (except Ref. (4)) it has been usual to call the characteristic impedance of the aerial $Z_0(x)$ twice the value given by (10).* We have therefore

$$2\zeta_0(x) = Z_0(x) = 60\left\{\sinh^{-1}\frac{l+x}{a} + \sinh^{-1}\frac{l-x}{a}\right\}. \tag{11}$$

---

* This is not always done, however. Cf. Ref. (15).

When $l \gg a$ we may write this in the approximate form

$$Z_0 = 120 \log_e \frac{2l}{a}, \tag{12}$$

which is very nearly the same as (11) except near the ends of the aerial. Z_0 is shown as a function of l/a in fig. 17.

The assumptions made above with regard to taking $i(x)$ outside the integral sign and neglecting retardation, while justified in the case of a transmission line, are by no means justified in the case of an aerial. The difference arises as follows. For a transmission line the contributions from the two conductors almost cancel except when x is nearly equal to x'. This is not so for an aerial, and although

$$(a) \; Z_0 = 120 \log_e \frac{2l}{a} \quad (b) \; Z_0 = 120 \left[\log_e \frac{2l}{a} - 1 \right]$$
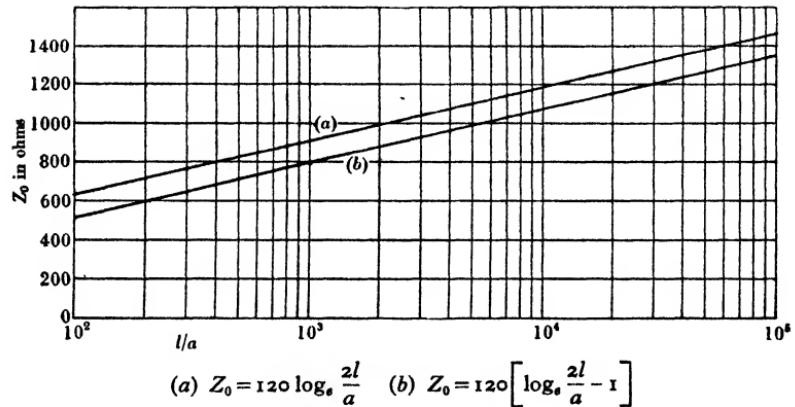
Fig. 17. Characteristic impedance of cylindrical aerial.

the integrand becomes very large when $x = x'$ if a is small, the integral only falls off logarithmically. This is the central weakness of the approximation assuming constant Z_0.

We have seen in § 3·4 how to calculate the real part of the input impedance of a half-wave aerial. As a first approximation to the reactive part of the impedance we might regard the aerial as a transmission line open-circuited at its free ends and write $Z = r + jx$, where

$$x = jZ_0 \cot \frac{2\pi l}{\lambda} \tag{13}$$

for values of l near $\frac{1}{4}\lambda$. This is not, however, a good approximation. It would indicate that the reactance of an aerial exactly $\frac{1}{2}\lambda$ long is zero, which is known to be untrue. However, we shall see that $\partial x/\partial l$ as calculated from (13) gives a very good approximation for the *variation* of x with l near $l = \frac{1}{4}\lambda$.

In order to get a reasonable approximation to the input impedance on the basis of the sinusoidal theory we proceed as follows(1, 2). Instead of calculating the flux of radiation over a large sphere as in §3·4 we evaluate the complex Poynting flux over the aerial itself. This should be zero except for the contribution from the source, but according to the sinusoidal theory we calculate it from the field produced by the sinusoidal current at the surface of the aerial. The Poynting flux outward from the surface of the aerial is

$$\iint [\mathbf{E} \times \mathbf{H}]\, dS, \tag{14}$$

taken over the surface of the aerial. Since in m.k.s. units H is equal in magnitude to the surface current density and at right angles to it, (14) becomes

$$-\int E i\, dx. \tag{15}$$

If Z is the impedance at the input we have, if i_0 is the current at the input,

$$Z i_0^2 = -\int E_s(x)\, i_s(x)\, dx, \tag{16}$$

where E_s is the field produced by the sinusoidal current i_s having a value i_0^* at the centre of the aerial.

The argument as we have given it is that generally employed and it will be appreciated that it is very unsound. It amounts to saying that the ratio of voltage to current at the input of the aerial under the actual conditions of transmitting is the same as for the very artificial distribution of driving force required to produce the sinusoidal current. There is no doubt, however, that equation (16) gives substantially correct results when a is so small that $\log_e 2l/a$ is large. Let us therefore try to see why this is so. The reason lies in the fact that $i_s(x)$ is a good approximation to the actual current when the aerial is used for transmitting. The reciprocity relation (42) of Chapter 3 may be used to show this as has been pointed out by Burgess(5), taking $E_s(x)$ as a field applied externally to the aerial. Since the tangential component of E must vanish this will produce in the aerial just exactly the sinusoidal current necessary to create the field $-E_s(x)$. We then have exactly, writing $e_r = Z_{11} i_0$,

$$Z_{11} i_0^2 = -\int E_s(x)\, i_t(x)\, dx, \tag{17}$$

where $i_t(x)$ is the correct form for the current when transmitting. When, however, $i_t(x)$ approximates closely to $i_s(x)$ (16) becomes a good approximation to (17).

The field $E_s(x)$ may now be evaluated using the vector potentials $\mathbf{A}(x)$ and $\phi(x)$. Let us consider first an aerial exactly half a wave-length long. The sinusoidal current will then be given by

$$i(x) = i_0 \cos \frac{2\pi x}{\lambda} = i_0 \cos kx, \qquad (18)$$

where $k = 2\pi/\lambda$.

For a thin cylindrical aerial this current may be taken as along the axis when evaluating $\mathbf{A}(x)$. This approximation has been examined by Hallén[6], who has shown that it is equivalent to having a slightly different radius for the aerial.

We then have from (6),

$$A_x(x') = \frac{\mu i_0}{4\pi} \int_{-l}^{l} \cos kx \, \frac{e^{-jkr}}{r} \, dx, \qquad (19)$$

where $r^2 = a^2 + (x - x')^2$. A_y and A_z are zero.

Equation (19) may be written in the form

$$A_x(x') = \frac{\mu i_0}{8\pi} \int_{-l}^{l} \left[\frac{e^{-jk(x+r)}}{r} + \frac{e^{-jk(r-x)}}{r} \right] dx. \qquad (20)$$

Also from the continuity relation between current and charge we have

$$\frac{\partial i}{\partial x} + \frac{\partial q}{\partial t} = 0, \qquad (21)$$

or

$$q(x) = -j i_0 / c \sin \frac{2\pi x}{\lambda}. \qquad (22)$$

Hence

$$\phi(x') = \frac{i_0}{8\pi c} \int_{-l}^{l} \left[\frac{e^{-jk(x+r)}}{r} - \frac{e^{-jk(r-x)}}{r} \right] dx. \qquad (23)$$

We therefore have

$$E_x(x') = -\frac{i_0}{8\pi ec} \int_{-l}^{l} \frac{\partial}{\partial x'} \left[\frac{e^{-jk(x+r)}}{r} - \frac{e^{-jk(r-x)}}{r} \right] dx$$

$$- \frac{j\omega \mu i_0}{8\pi} \int_{-l}^{l} \left[\frac{e^{-jk(x+r)}}{r} + \frac{e^{-jk(r-x)}}{r} \right] dx. \qquad (24)$$

Now using the fact that $\partial r/\partial x = -\partial r/\partial x'$, the two integrals combine into the simple form

$$E_x(x') = -\frac{i_0}{8\pi ec} \int_{-l}^{l} \frac{\partial}{\partial x} \left[\frac{e^{-jk(x+r)}}{r} - \frac{e^{-jk(r-x)}}{r} \right] dx.$$

Using the relations $1/c = (\epsilon\mu)^{\frac{1}{2}}$, $(\mu\epsilon)^{\frac{1}{2}} = 120\pi$ and $kl = \frac{1}{2}\pi$, we have, on putting in the limits of integration,

$$E_x(x) = -j \cdot 30i_0\left[\frac{e^{-jkr_1}}{r_1} + \frac{e^{-jkr_2}}{r_2}\right], \tag{25}$$

where $r_1^2 = a^2 + (x-l)^2$, $r_2^2 = a^2 + (x+l)^2$.

From equation (17) the impedance Z_{11} is therefore seen to be given by

$$Z_{11} = 30j\int_{-l}^{l}\cos kx\left[\frac{e^{-jkr_1}}{r_1} + \frac{e^{-jkr_2}}{r_2}\right]dx. \tag{26}$$

For a thin wire this may be written

$$Z = 60j\int_{0}^{l}\cos kx\left[\frac{e^{-jk(l-x)}}{l-x} + \frac{e^{-jk(l+x)}}{l+x}\right]dx \tag{27}$$

$$= 30\int_{0}^{l}\left[\frac{1 + e^{+2jkx}}{l-x} + \frac{1 + e^{-2jkx}}{l+x}\right]dx. \tag{28}$$

Putting $y = 2k(l-x)$ in the first integral and $y = 2k(l+x)$ in the second, we have, since $l = \frac{1}{4}\lambda$,

$$Z = 30\int_{0}^{2\pi}\left(\frac{1 - e^{-jy}}{y}\right)dy \tag{29}$$

$$= 30\int_{0}^{2\pi}\frac{1 - \cos y}{y}\,dy + 30j\int_{0}^{2\pi}\frac{\sin y}{y}\,dy. \tag{30}$$

These integrals may be evaluated numerically or may be expressed in terms of the Si and Ci functions.

We have

$$Z = 30[\{\gamma + \log_e(2\pi) - \text{Ci}(2\pi)\} + j\text{Si}(2\pi)], \tag{31}$$

where γ is Euler's constant, $0.5772\ldots$

The numerical value of (30) or (31) is

$$Z = r_{11} + jx_{11} = 73.2 + j.42.5 \text{ ohms.} \tag{32}$$

The above calculation is only possible when the length of the aerial is $\frac{1}{2}\lambda$ (or an odd multiple of $\frac{1}{2}\lambda$), otherwise the integral (27) diverges as $a \to 0$. If we have a current of the more general form (1) the principal contribution to the integral arising from the region where r_1 or r_2 become very small may be shown in a similar manner to that used to obtain (11) to be approximately

$$-jZ_0\cot kx + \text{terms which do not tend to } \infty \text{ as } a \to 0.$$

When l is nearly equal to $\frac{1}{4}\lambda$ the principal term does not differ much from (13) obtained on the analogy of transmission-line theory. For a thin aerial whose length $2l$ is approximately $\frac{1}{2}\lambda$ we may write the input impedance in the approximate form

$$Z = 73 \cdot 2 + j \cdot 42 \cdot 5 - jZ_0 \cot \frac{2\pi l}{\lambda}, \tag{33}$$

or since this applies only when $l \simeq \frac{1}{4}\lambda$

$$Z = 73 \cdot 2 + j \cdot 42 \cdot 5 + j \frac{2\pi Z_0}{\lambda} \Delta l, \tag{34}$$

where $2\Delta l$ is the excess in total length of the aerial over half a wave-length and Z_0 is given by (12). The formula (34) does not take into account variation of the resistive and reactive parts of the impedance of a half-wave aerial with a and is only valid to an accuracy in which $1 \Big/ \left(2 \log_e \dfrac{2l}{a} \right)$ is neglected compared with unity, as shown by a second approximation.

We shall later give a more accurate formula, but for thin aerials (34) gives values in good agreement with experiment for values of Δl of a few per cent.

It will be seen that the percentage change in reactance increases as Z_0 increases, i.e. it is greatest for thin aerials. The length of a thin resonant half-wave dipole may be obtained from (34). It is given by

$$\frac{\lambda}{2} \left[1 - \frac{60 \operatorname{Si}(2\pi)}{\pi Z_0} \right] \tag{35}$$

or

$$\frac{\lambda}{2} \left[1 - \frac{27}{Z_0} \right]. \tag{36}$$

This is shown as a function of $\lambda/4a$ in fig. 19, curve (a).

When we attempt to apply the above method to a full-wave dipole we experience a grave difficulty. The input impedance turns out to be infinite. This is not surprising, as, on the sinusoidal theory, the current at the input is zero. We may, however, obtain a fairly good approximation to the input impedance as follows, by applying transmission-line theory, assuming a constant value for the characteristic impedance. Let us consider a full-wave dipole as two half-wave dipoles in line, spaced $\frac{1}{2}\lambda$ apart. We may calculate by means of mutual impedances (see §§ 2·1 and 4·9) the centre impedance of each half. Let i_0 be the current at the centre of each. Then by ordinary transmission-line theory the voltage at a point $\frac{1}{4}\lambda$ away

will be $i_0 Z_0'$, where $Z_0' = 120 \log_e l/a$. The length of each half-wave dipole is l. The complex power will be given by the following alternative expressions which may be equated

$$\frac{i_0^2 Z_0'^2}{Z} = 2i_0^2(Z_{11} + Z_{12}) \tag{37}$$

or

$$Z = \frac{Z_0'^2}{2[Z_{11} + Z_{12}]}, \tag{38}$$

where Z_{11} and Z_{12} are the self- and mutual impedances of the equivalent half-wave dipoles, and $Z_0' = Z_0 - 120 \log_e 2$.

As a first approximation we use the values of Z_{11} and Z_{12} for infinitely thin dipoles, namely,

$$Z_{11} = 73 \cdot 2 + j \cdot 42 \cdot 5 = r_{11} + jx_{11}, \quad Z_{12} = 25 \cdot 1 + j \cdot 19 \cdot 8 = r_{12} + jx_{12}.$$

From equation (31) and equation (85) of §4·9 we note that

$$x_{11} + x_{12} = 15[4\operatorname{Si}(2\pi) - \operatorname{Si}(4\pi)]. \tag{39}$$

It is more convenient in the case of full-wave dipoles to work in terms of admittance rather than impedance. If Y_λ is the admittance of a dipole exactly λ long we have from (37)

$$Z_0'^2 Y_\lambda = 2(Z_{11} + Z_{12}) \tag{40}$$

or

$$Z_0'^2 Y_\lambda = 196 \cdot 6 + j \cdot 124 \cdot 6 \text{ ohms.} \tag{41}$$

The first approximation to the admittance Y_{11} of a full-wave aerial is then obtained by adding to (38) the admittance of a line of length l of characteristic impedance Z_0. We then have

$$Z_0'^2 Y_{11} = 196 \cdot 6 + j \cdot 124 \cdot 6 + jZ_0 \tan \frac{2\pi l}{\lambda}, \tag{42}$$

or since this is only valid for $l \simeq \tfrac{1}{2}\lambda$

$$Z_0'^2 Y_{11} = 196 \cdot 6 + j \cdot 124 \cdot 6 + 2\pi Z_0 \frac{\Delta l}{\lambda}. \tag{43}$$

The anti-resonant resistance Z_r is given approximately by

$$Z_r = Z_0'^2/196 \cdot 6 \text{ ohms.} \tag{44}$$

The formula (43) is the counterpart for full-wave aerials of (34). We see from (39) and (42) that the length of a resonant full-wave aerial is

$$\lambda \left\{ 1 - \frac{30[4\operatorname{Si}(2\pi) - \operatorname{Si}(4\pi)]}{\pi Z_0} \right\} \tag{45}$$

or

$$\lambda \left[1 - \frac{40}{Z_0} \right].$$

This is shown as a function of $\lambda/2a$ in fig. 20, curve (a). It will be seen that the impedance of a full-wave dipole increases with Z_0 and tends to infinity as a tends to zero. The value of the resistive part of the impedance varies with l in the second approximation, but as it is near a maximum this variation is not great and the value (44) for the resonant impedance is a fair approximation in the case of thin aerials. This is shown as a function of $\lambda/2a$ in fig. 22, curve (a).

4·2. The fundamental integral equation of aerial theory

Before proceeding further we shall derive an integral equation which, if we could solve it, would lead to an exact solution of the aerial problem. It is the starting-point of several attempts to find approximate solutions, using the idea of a constant characteristic impedance as a zero-order approximation, giving results equivalent to those obtained by the sinusoidal theory, and using this approximation to obtain higher order approximations.

The fundamental integral equation we have already written down in the form (2). For a thin cylindrical aerial it takes the form

$$A(x') = \frac{\mu}{4\pi} \int_{-l}^{l} K(x - x') \, i(x) \, dx, \qquad (46)$$

where
$$K(x) = \exp\left\{ -jk(x^2 + a^2)^{\frac{1}{2}} \right\} / (x^2 + a^2)^{\frac{1}{2}}, \qquad (47)$$

making the approximation that the surface current can be replaced by an axial current. If we could determine $A(x)$, (46) would be an integral equation to determine $i(x)$, the current distribution. Fortunately, there is a very important result which simplifies matters considerably. It is that the potentials $\mathbf{A}(x)$ and $\phi(x)$ are accurately sinusoidal in form at the surface of the aerial. This is most readily proved by introducing the Hertzian vector $\mathbf{\Pi}$ related to $\mathbf{A}$ and ϕ by the equations (Ref. (1) of Chapter 3, p. 28)

$$\mathbf{A} = \frac{1}{c^2} \frac{\partial \mathbf{\Pi}}{\partial t}, \quad \phi = -\operatorname{div} \mathbf{\Pi}. \qquad (48)$$

In terms of $\mathbf{\Pi}$ we have for a linear aerial, since only Π_x is not zero,

$$E_x = \frac{\partial^2 \Pi_x}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 \Pi_x}{\partial t^2}, \qquad (49)$$

since $E_x = 0$ at the surface of the aerial Π_x satisfies the wave equation and is sinusoidal. From (48) it follows that $\mathbf{A}$ and ϕ are also sinusoidal.

The various approximate methods of solving (46) proceed essentially by assuming a particular sinusoidal form for $A(x)$ and determining the constants in it so as to satisfy the boundary conditions. Before examining these, however, we shall use (46) to prove the very important result that the distribution of current in a very thin aerial is sinusoidal and incidentally introduce more elegantly the concept of 'characteristic' or 'wave' impedance.

4·3. Proof that the current in a very thin aerial is sinusoidal

Since, when a is small, the main contribution to the integral on the right-hand side of (46) comes from values of x near x', let us write (46) in the form

$$A(x') = \frac{\mu}{4\pi} i(x') e^{-jka} \int \frac{dx}{\{(x-x')^2 + a^2\}^{\frac{1}{2}}}$$
$$+ \frac{\mu}{4\pi} \int \frac{i(x) e^{-jk[(x-x')^2 + a^2]^{\frac{1}{2}}} - i(x') e^{-jka}}{\{(x-x')^2 + a^2\}^{\frac{1}{2}}} dx. \tag{50}$$

The first integral tends to infinity as a tends to zero. For small values of a it is equal to

$$Z_0 i(x')/2c, \tag{51}$$

where Z_0 is given by (11). As a tends to zero the second integral tends to

$$\frac{\mu}{4\pi} \int \frac{i(x) e^{-jk(x-x')} - i(x')}{x - x'} dx.$$

On expanding the integral in the region round $x = x'$ it is clear that the integral converges. We therefore have

$$cA(x) = \tfrac{1}{2} Z_0 i(x) + c\chi(x), \tag{52}$$

where $\chi(x)$ remains finite as $a \to 0$. As $a \to 0$ we have

$$i(x') \to \text{const.} \times A(x').$$

Now we have proved in §4·2 that $A(x)$ is sinusoidal. Therefore $i(x)$ tends to sinusoidal form as $a \to 0$. Without much more complicated analysis we cannot determine the extent to which $i(x)$ departs from sinusoidal form for a given value of a.

4·4. Wave impedance

The form (52) suggests a new definition of Z_0 as the ratio of current to the vector potential multiplied by a constant, i.e.

$$Z_0(x) = 2cA(x)/i(x). \tag{53}$$

This method of defining Z_0 avoids the idea of inductance and capacity per unit length and gives the usual value when applied to a transmission line. We see, however, from (52) that defined this way

$$Z_0(x) = 120 \log_e \frac{2l}{a} + \frac{2\chi(x)}{i(x)}, \tag{54}$$

and is not constant in the second approximation. 'Wave impedance' is therefore a better name than 'Characteristic impedance'.

4·5. Approximate methods of solution of the integral equation

Various approximate methods of solution have been attempted. These normally consist of assuming effectively a sinusoidal form for $A(x)$ and taking as zero-order approximation for $i(x)$ the form $2cA(x)/Z_0$. Various writers have chosen different starting values for Z_0. Hallén[6] has used the form (12) but has tried to correct it near the ends of the aerial. Marion Gray[17] and Middleton and King[19] have taken a considerably more complex form as a zero-order approximation. The constants of $A(x)$ are chosen to fit the boundary conditions as applied to $i(x)$.

4·5·1. Hallén's method

The first-order approximation is obtained by taking

$$i(x) = 2cA(x) Z_0 + \frac{1}{\Omega} j(x), \tag{55}$$

where $\Omega = 2 \log_e 2l/a$ and solving approximately for $j(x)$ by means of the integral equation. The work of Hallén has been extended numerically by King and Blake[7], who give a great variety of curves for the impedance of aerials having all lengths up to a few wavelengths. In Hallén's method the input impedance is given in the form

$$Z_{11} = -j \frac{Z_0 \cos(2\pi l/\lambda) - \alpha(l) + j\beta(l) + \zeta(l) Z_0^{-1}}{\sin(2\pi l/\lambda) + \gamma(l) Z_0^{-1} - j\delta(l) Z_0^{-1} + \xi(l) Z_0^{-2}}, \tag{56}$$

where, as before, $Z_0 = 120 \log_e(2l/a)$.

The functions α, β, γ, δ are real functions of l and are given as analytical expressions involving Si and Ci functions. The functions ζ and ξ are complex and have to be calculated by numerical methods. The labour involved in this 'third approximation' is very great and has been carried out to some extent by Scott* and Bowkamp[18].

---

* Unpublished calculations by J. M. C. Scott performed for Ministry of Supply.

We shall call the approximation neglecting ζ and ξ the 'second approximation', as it gives correction terms of the order of $1/Z_0$ to the 'first approximations' given by (34) and (43). The second approximation is then

$$Z_{11} = -j\frac{Z_0\cos(2\pi l/\lambda) - \alpha(l) + j\beta(l)}{\sin(2\pi l/\lambda) + \gamma(l)Z_0^{-1} - j\delta(l)Z_0^{-1}}. \tag{57}$$

In the limit as a tends to zero $\beta(l)$ tends to r_{11} and $\alpha(l)$ to x_{11}, the resistance and reactance of a half-wave aerial. As a tends to zero, therefore (56) reduces to the form (34) for values of l near $\frac{1}{4}\lambda$. To a higher approximation, for values of l near $\frac{1}{4}\lambda$, (56) may be written in the form

$$Z_{11} = r_{11} + jx_{11} + j\frac{2\pi Z_0\Delta l}{\lambda} + \frac{d\Delta l}{\lambda}(1-j) - \frac{1}{Z_0}(b+jc), \tag{58}$$

where b, c, d are constants. The numerical values of b, c, d are as follows: $b = 5400$, $c = 9700$, $d = 800$. (58) may then be regarded as the second approximation to the impedance of a half-wave aerial.

The impedance of an aerial exactly $\frac{1}{2}\lambda$ long is given by

$$Z_{\frac{1}{2}\lambda} = \left(73\cdot2 - \frac{5400}{Z_0}\right) + j\left(42\cdot5 - \frac{9700}{Z_0}\right) \text{ ohms.} \tag{59}$$

The input resistance at resonance is given by

$$r_{\text{res.}} = \left(73\cdot2 - \frac{10,500}{Z_0}\right) \text{ ohms,} \tag{60}$$

and is shown as a function of $\lambda/4a$ in fig. 21, curve (b). The change in reactance for a change Δl in l is given by

$$Z = j\frac{2\pi Z_0}{\lambda}\left(1 - \frac{100}{Z_0}\right)\Delta l. \tag{61}$$

The length for zero reactance is given by

$$l = \frac{\lambda}{4}\left[1 - \frac{27}{Z_0} + \frac{2300}{Z_0^2}\right]. \tag{62}$$

This is shown as a function of $\lambda/4a$ in fig. 19, curve (c). The formulae (58), (59) and (62) give the corrections to the order of $1/Z_0$ to the formulae (34), (32) and (36) obtained from the simple sinusoidal theory.

When l is approximately equal to $\frac{1}{2}\lambda$, $\sin(2\pi l/\lambda)$ is small, and we must approximate more carefully to (56). When $l = \frac{1}{2}\lambda$ the input impedance is given by

$$Z = \frac{Z_0^2}{\delta(l) + j\gamma(l)}, \tag{63}$$

which is in the same form as (38).

The variation of the resistance near the resonant point is so rapid that formula (57) must be used for its calculation. However, the maximum value is very nearly equal to the resonant value. The exact curves calculated by King and Blake using (55) indicate, in fact, that the maximum resistance is given approximately by (44), and the resonant length of a full-wave dipole by (45). The second approximation to (44) is given by the expressions

$$Z_r = \frac{Z_0'^2}{196\cdot6} + \frac{Z_0'}{2\cdot5} \tag{64}$$

or

$$= \frac{Z_0^2}{196\cdot6} - \frac{Z_0}{2\cdot3}. \tag{65}$$

This is shown as a function of $\lambda/2a$ in fig. 22, curve (b).

The results of these calculations do not agree very well with experiment, as will be seen from figs. 19–24. In particular, this method gives too high values for the resonant impedance of full-wave aerials. This is not surprising in view of the nature of the approximation and the lack of knowledge of the rate of convergence of successive approximations. It is found from the work of Bowkamp[18] that the convergence is not rapid and the results are still in some disagreement with experiment, even when a further stage of approximation is undertaken. Schelkunoff[16] has shown that Hallén's approximation involves the tacit assumption in the evaluation of certain integrals that $l \ll \lambda$. This has been avoided by Marion C. Gray[17] by taking a much more complicated expression than $\frac{1}{2}\log_e 2l/a$ as the expansion parameter. By this means a more rapid rate of convergence is obtained, and the results of the calculation are in much better agreement with experiment. Another approach has been made by Middleton and King[19], who have also succeeded in obtaining an approximation in good agreement with experiment, the most interesting new point in their results being the higher values, about 70 ohms, for the resonant resistance of

a half-wave dipole. These writers give an interesting comparison of the various methods used so far for approximating to the solution of Hallén's integral equation.

Another interesting method of approximating to the solution of Hallén's integral equation was used by H. G. Booker.* The value of $Z(x)$ was chosen in the first approximation so as to give the *same* value in the second approximation, the integral equation having been put in a form suitable for successive approximations by means of a recurrence relation between the nth and the $(n+1)$th forms for $A(x)$. Four approximations were required to give a reasonable value for the impedance of a $\frac{1}{2}\lambda$ dipole. This value (for $\lambda/a = 800$) was $81 + j \cdot 44$ ohms. Further approximations, however, showed that the convergence was slow, the reactive part failing to tend to a limit.

4·6. Approach from conical aerials

A completely different method of approach to the problem of calculating theoretically the various quantities for thin aerials of arbitrary shape has been made by Schelkunoff(4). This starts by dealing with the problem of biconical aerials which may be treated by exact mathematical methods and then generalizing the analysis by an approximate method to deal with thin aerials of arbitrary shape.

The biconical aerial is particularly simple since exact solutions of the wave equation in spherical co-ordinates may be found for the space near the aerial, and these may be fitted to solutions in free space outside the 'aerial region' defined as a sphere with centre at the apex of the cones and passing through the ends of the aerial (fig. 18).

One of the solutions in the 'aerial region' corresponds to propagation of the principal mode along the uniform transmission line consisting of the two cones. That these form a uniform transmission line may be more readily seen by considering cones whose semi-angle is nearly equal to 90°. Both the inductance and capacity per unit length are independent of r, the distance from the apex, as the length of the circular lines of electric force and the radius of the cone are both proportional to r. In fact, if ϕ is the semi-angle of the cone,

$$L = \frac{\mu}{\pi}\log \cot \tfrac{1}{2}\phi, \quad C = \pi\epsilon/\log \cot \tfrac{1}{2}\phi, \tag{66}$$

* Unpublished calculations carried out for Ministry of Aircraft Production in 1941.

giving $$Z_0 = (L/C)^{\frac{1}{2}} = 120 \log_e (\cot \tfrac{1}{2}\phi).$$ (67)

When ϕ is small $$Z_0 = 120 \log_e \left(\frac{2l}{a}\right),$$ (68)

where l is the length of each cone and a the radius of the base. Note that there is no approximation in defining L and C as for a cylindrical aerial. They may be deduced precisely from the solution of Maxwell's equations for propagation between two cones. The transmission-line propagation corresponds, however, only to the 'principal mode'. Superposed on this are secondary modes (as in wave-guide theory). The mathematical method consists of fitting the solution consisting of principal mode and secondary modes for the 'aerial region' to a solution representing outgoing waves in the region outside.
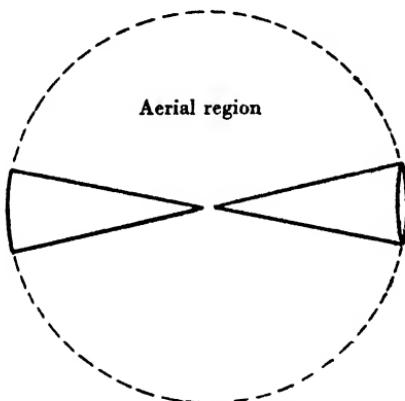
Fig. 18. Biconical aerial.

It is shown that the voltage $V(r)$ at any point along the cone is just equal to $V_0(r)$, that due to the principal mode. This corresponds to a result that we have had already that the potentials are propagated sinusoidally. The current $i(r)$ consists of the current $i_0(r)$ of the principal mode, plus that due to the secondary modes, i.e.

$$i(r) = i_0(r) + i_1(r) + i_2(r).$$ (69)

$i_0(r)$ is, of course, sinusoidal and the contributions from the secondary modes give the deviation from sinusoidal form.

Reflexion takes place at the ends of the cones, but some of the energy goes off into space as radiation. It is shown that this is

equivalent to terminating the transmission line with a certain impedance Z_t, given by

$$Z_t = Z_0^2/[G(L) + jF(L)], \qquad (70)$$

where $L = 2\pi l/\lambda$.

Note that $G(L) + jF(L)$ is just the impedance which would be seen $\frac{1}{4}\lambda$ from the ends of the antenna, and corresponds to the input impedance for a $\frac{1}{2}\lambda$ aerial. The input impedance in general is then given by

$$Z_t = Z_0 \frac{(G+jF)\cos(L - \frac{1}{2}\pi) + jZ_0 \sin(L - \frac{1}{2}\pi)}{Z_0 \cos(L - \frac{1}{2}\pi) + j(G+jF)\sin(L - \frac{1}{2}\pi)}. \qquad (71)$$

The functions $G(L)$, $F(L)$ are expressible in terms of Si and Ci functions. This is very satisfactory as far as conical aerials are concerned.

For a thin conical aerial exactly half a wave long $G(L)$ is just equal to the real part of (31). $F(L)$ is, however, equal to

$$30\,\mathrm{Si}(2\pi) + 60\,\mathrm{Si}(\pi)$$

which has the value 152·5 ohms. Thus, as we should expect, the radiation resistance of a thin conical aerial is the same as for a thin cylindrical one. The reactance, however, is much greater. The resonant length for a conical half-wave aerial is

$$\frac{\lambda}{2}\left[1 - \frac{98}{Z_0}\right], \qquad (72)$$

with Z_0 given by (68). This should be compared with the value (36) for a cylindrical aerial.

Having solved successfully the problem of a conical aerial Schelkunoff then adapted the method to give an approximate solution for aerials of other shapes. The starting-point of the approximation is to define a function which is the analogue of the characteristic impedance of the conical aerial, which as we have seen is a constant. Let ρ be the radius of the aerial at distance r from the origin. Then *define* $Z(r, \rho)$ by means of the equation

$$Z(r, \rho) = 120 \log_e \frac{2r}{\rho}. \qquad (73)$$

This is written down on the analogy of the conical aerial, but also turns up in the approximate solution of the problem of the propagation of the principal mode down a transmission line consisting of

wires of varying section. $Z(r, \rho)$ may be regarded as the characteristic impedance of the line and varies with distance r since, of course, ρ is a function of r. We now proceed to define an average characteristic impedance Z_0 of the aerial by

$$Z_0 = \frac{1}{l} \int_0^l Z(r, \rho) \, dr. \tag{74}$$

For example, for a cylindrical aerial,

$$Z_0 = 120 \left(\log_e \frac{2l}{a} - 1 \right). \tag{75}$$

For a spheroidal aerial $\quad Z_0 = 120 \log_e \frac{l}{a}, \tag{76}$

where $2a$ is the length of the minor axis and $2l$ that of the major axis.

△ Smith and Holt Smith ($S = \frac{1}{2}$ in.)　　○ Essen and Olliver
× Rösseler, Vilbig and Vogt　◇ Brown and Woodward　▽ Cochrane
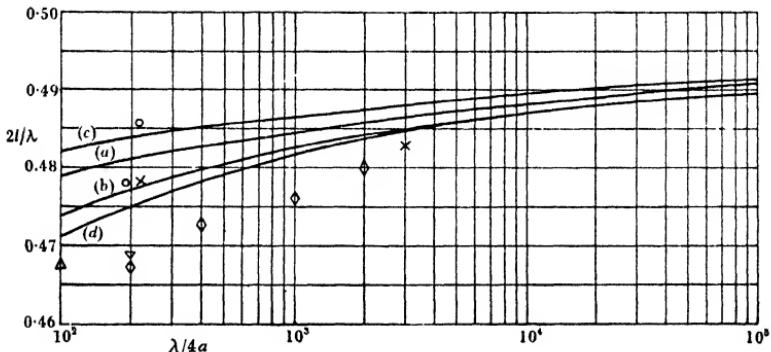
Fig. 19. Resonant length of cylindrical half-wave dipole of radius a. Curve (a), simple sinusoidal theory. Curve (b), sinusoidal theory with Schelkunoff's value of Z_0. Curve (c), King and Blake. Curve (d), Schelkunoff.

The form (75) should be compared with (12). This form was in fact used by some early writers who worked on the sinusoidal theory and is undoubtedly a better approximation than (12). It is also shown as a function of l/a in fig. 17, curve (b).

We now proceed to solve the problem of the aerial with non-uniform characteristic impedance by using (73) and the theory of propagation down a non-uniform line to derive the principal mode in the 'aerial region'. For the secondary modes we use the values obtained for a conical aerial, but having Z_0 given by (74). Since

the contribution from the secondary modes is in the nature of a first-order correction this seems a reasonable procedure.

One interesting check on this theory is that it leads to the result that, to the first order in $1/Z_0$, the reactance of a spheroidal aerial

○ Essen and Olliver   ✳ D. D. King   ◇ Brown and Woodward
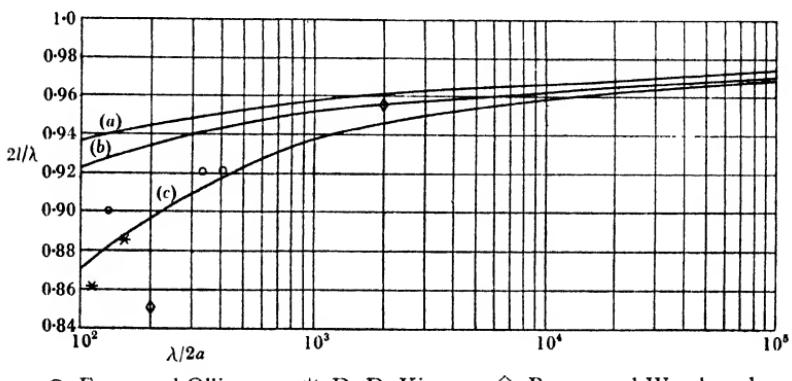
Fig. 20. Resonant length of cylindrical full-wave dipole. Curve (a), simple sinusoidal theory. Curve (b), sinusoidal theory with Schelkunoff's value of Z_0. Curve (c), Schelkunoff.

△ Smith and Holt Smith ($S = \frac{1}{2}$ in.)   ○ Essen and Olliver
× Rösseler, Vilbig and Vogt   ◇ Brown and Woodward   ▽ Cochrane
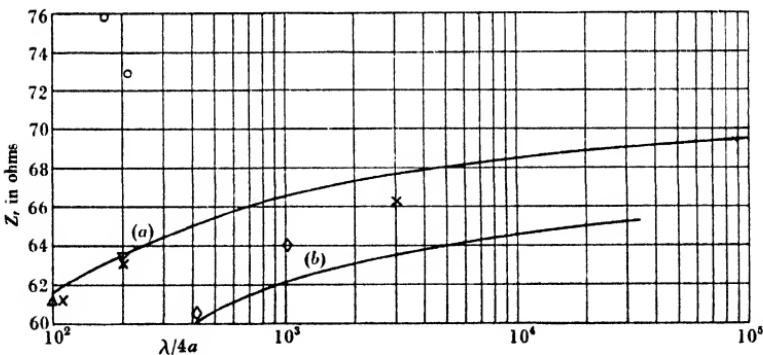
Fig. 21. Resonant resistance of half-wave dipole. Curve (a), Schelkunoff. Curve (b), King and Blake.

is zero; this result is also obtained from a more rigorous solution of the problem as we shall see in §4·7. Moreover, it also shows up how difficult it is to decide what is 'an equivalent cylindrical aerial' to correspond to a spheroidal aerial, and so make use of the exact solution available for the latter.

For a cylindrical aerial of radius a and length exactly $\frac{1}{2}\lambda$ the input impedance takes the simple form

$$Z_{\frac{1}{2}\lambda} = \frac{Z_0}{Z_0 - 21}\,[r_0 + jx_0], \tag{77}$$

where $Z_0 = 120\,[\log_e \lambda/2a - 1]$, and r_0, x_0 are the values of the resistance and reactance for an infinitely thin aerial, $73 \cdot 2$ and $42 \cdot 5$ ohms respectively. Equation (77) may be written approximately as

$$Z_{\frac{1}{2}\lambda} = \left(73 \cdot 2 + \frac{1540}{Z_0}\right) + j\left(42 \cdot 5 + \frac{890}{Z_0}\right). \tag{78}$$

This should be compared with (59), the form given by Hallén's approximation.

● Various points quoted by Schelkunoff   ✻ D. D. King
○ Essen and Olliver   ◇ Brown and Woodward
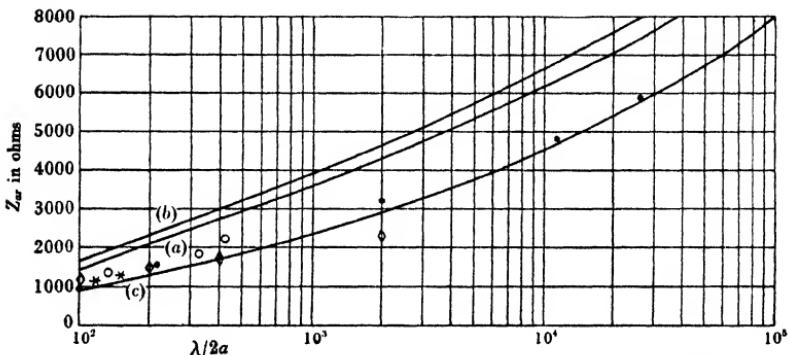
Fig. 22. Anti-resonant resistance of full-wave dipole. Curve (a), simple sinusoidal theory. Curve (b), King and Blake. Curve (c), Schelkunoff and also sinusoidal theory with Schelkunoff's value of Z_0.

The rate of change of the input resistance with length for an aerial nearly half a wave long is found to be independent of Z_0 to the order of $1/Z_0$ and equal to $2 \cdot 5$ ohms for a $1\,\%$ change. The rate of change of reactance is as given by (13) with Schelkunoff's value of Z_0. For $l \simeq \frac{1}{4}\lambda$ we then have for the input impedance

$$Z = Z_{\frac{1}{2}\lambda} + 1000\frac{\Delta l}{\lambda} + j\,\frac{2\pi Z_0 \Delta l}{\lambda}. \tag{79}$$

The values obtained for the resonant resistance are given in fig. 21 and those for the resonant length in fig. 19. It will be seen that these do not differ much from what is given by the simple formula (36) provided Schelkunoff's value of Z_0 is used.

For a full-wave dipole Schelkunoff gives a more complicated formula. In the range of l/a from 10^2 to 10^5, however, the numerical values as given by his formula do not differ appreciably from those given by the simple formula (45) provided Schelkunoff's value of

$\triangle$ Smith and Holt Smith ($S = \frac{1}{2}$ in.)   × Rösseler, Vilbig and Vogt
$\diamondsuit$ Brown and Woodward $\triangledown$ Cochrane
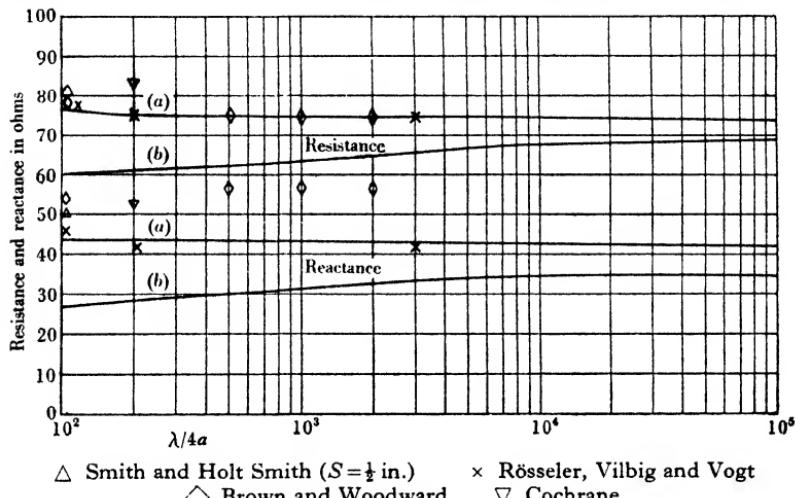
Fig. 23. Input impedance of cylindrical dipole exactly $\frac{1}{2}\lambda$ long. Curve (a), Schelkunoff. Curve (b), King and Blake.

$\triangle$ Smith and Holt Smith $\bigcirc$ Essen and Olliver
$*$ Rösseler, Vilbig and Vogt. $\diamondsuit$ Brown and Woodward $\triangledown$ Cochrane
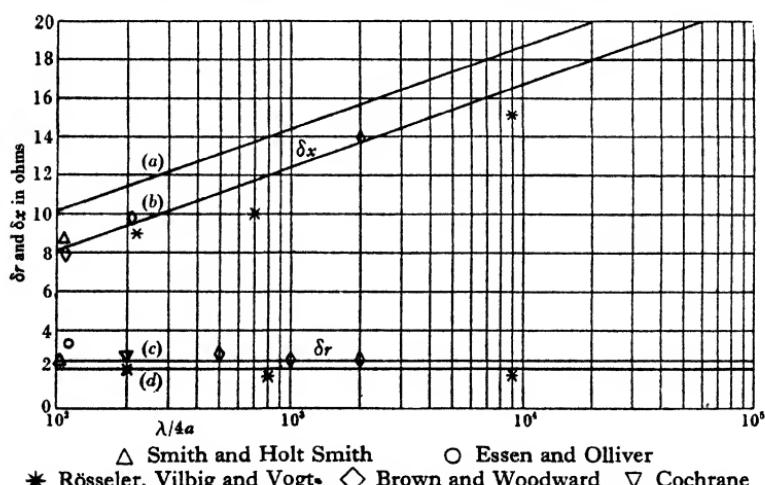
Fig. 24. Rate of change of reactance and resistance of half-wave dipole with length. Curve (a), δx—simple sinusoidal theory. Curve (b), δx—sinusoidal theory with Schelkunoff's value of Z_0. Curve (c), δr—Schelkunoff. Curve (d), δr—King and Blake. N.B. King and Blake's values and Schelkunoff's values for δx do not differ significantly from Curve (b).

Z_0 is used. The values of resonant length to the first order agree with (44), but an appreciable correction is given by the second-order terms as shown in fig. 20. This brings the theoretical values into better agreement with experiment. Apart from the usual disagreement between the theoretical and measured values of resonant resistance of a half-wave dipole due to effect of feeding arrangements the theoretical values given by Schelkunoff agree very well with the available experimental data (see §4·8 and figs. 19–24).

Schelkunoff's method enables the current in the aerial to be calculated fairly readily. Curves are given in his paper for a full-wave conical aerial with $Z_0 = 1000$ ohms showing how the current is non-zero at the input, and separating the current due to the principal mode and the phase and quadrature components due to the secondary modes. Only the first few modes contribute appreciably. As $Z_0 \to \infty$ the secondary current becomes vanishingly small, leaving only the 'principal' current which is sinusoidal.

4·7. Approach from spheroidal aerials

The spheroid is another geometrical form which lends itself to exact treatment. The wave equation is 'separable' in spheroidal co-ordinates and exact solutions are possible which in theory may be fitted to the boundary conditions, for example, those defining the driving conditions in the case of a driven aerial.

Abraham[8] has dealt with the free oscillations of a spheroid and has thus determined the resonant lengths of spheroidal aerials. His analysis has been extended by Page and Adams[9] to deal with a spheroidal aerial excited by a uniform field and by Chu and Stratton[10] to deal with an aerial driven at the centre. The work of Chu and Stratton is the most extensive as regards the determination of the input impedance. An interesting result found by all these writers is that if ϵ is the ratio of major to minor axis, then to the first order in $1/\log_e \epsilon$ the first resonance occurs when $\lambda = 4l$, where $2l$ is the length of the major axis. This result as we have seen was also found by Schelkunoff. Abraham's value is

$$\lambda = 4l \left[1 + \frac{5040}{(120 \log_e \epsilon + 83)^2} \right]. \tag{80}$$

In considering the driving-point impedance Chu and Stratton ran into the usual trouble arising from the gap at the centre. They found that the value of the reactance fails to converge as the width

of the gap tends to zero. They evaded the difficulty, however, by noting that the contribution to the impedance can be separated into that arising from the various resonances, and that only the first few contribute appreciably when the aerial is short unless the gap is very small. By neglecting the contribution from the higher modes the convergence difficulty is by-passed. This method is very elegant, but it is doubtful if the results are of much practical value for application to other forms of aerials. It was hoped that a cylindrical aerial with radius of the same order as the minor axis of the spheroid would behave in a similar manner. The work of Schelkunoff has, however, shown clearly that this is not so, and it is rather difficult to choose an 'equivalent spheroid' for a cylindrical aerial.

(a) $l/a = 700$ (b) $l/a = 7$.
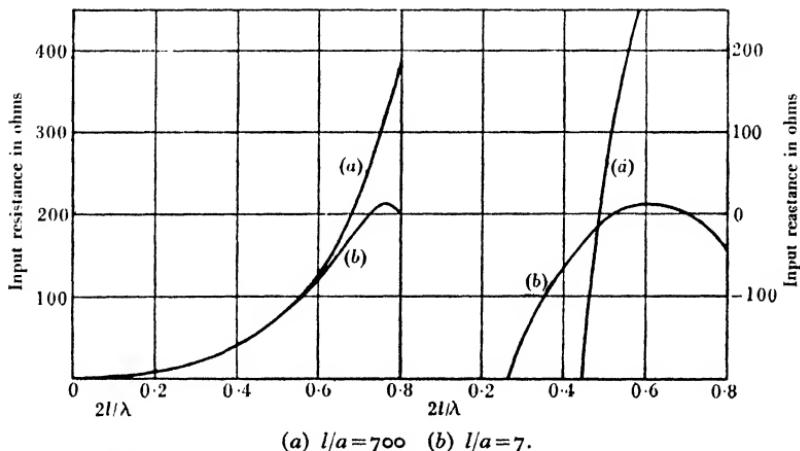
Fig. 25. Input resistance and reactance of spheroidal aerials.

The chief importance of this form of treatment, however, lies in the fact that it is the only method which gives a valid treatment of the problem when l/a is not large. In fact, the method has been applied by Chu and Stratton(11) to the case of a sphere when $l = a$. By this means the qualitative behaviour of very fat aerials used for wide-band working may be discussed. For example, in fig. 25, we compare the behaviour of two spheroidal aerials having l/a equal to 7 and 700. It will be seen that the fat aerial shows a very small variation of reactance with l. This is the property which is utilized for wide-band working. This trend has, of course, been clear from the various treatments of thin aerials where we have seen that the variation of reactance with length is proportional to $1/Z_0$. It is,

however, instructive to follow the trend when l/a is no longer large and the approximate treatments fail completely.

4·8. Comparison of theory and experiment

A good deal of experimental work has been carried out from time to time on the measurement of the characteristics of aerials. Until recently, however, when the accurate measurement of impedances became possible at very high frequencies, most of the work is open to question, as far as comparison with theory is concerned, though, of course, of practical importance from a design point of view. Measurements at wave-lengths greater than a few metres cannot be conveniently made on dipoles and have been made on unipoles with one end near an earth plane. Unless considerable precautions are taken to provide an earth plane of high conductivity losses incurred in it may vitiate the results.

Measurements on dipoles have been, in fact, extremely scanty. Apart from those reported by Smith and Holt Smith and by Cochrane* (Ref. (3), Chapter 1), the only other recent measurements are those of Essen and Olliver (Ref. (4), Chapter 1). The former measurements were made with wave-lengths of 12 m. and 6 m. using an h.f. impedance bridge, in the course of the design of radar receiving aerials, and unfortunately only refer to half-wave aerials having two ratios (200 to 1 and 400 to 1) of wave-length to diameter. Since, however, the behaviour near the resonant point was carefully studied, and accurate values obtained for the resonant length and rate of change of resistance and reactance, these experiments provide a useful check of the theory. Moreover, these experiments give the only available information on the effect on the input impedance of varying the gap at the centre of the dipole. The results of the experiments are summarized in fig. 3, Chapter 1, and are plotted with the theoretical results in figs. 19, 21 and 24. The rate of change of resistance and reactance are in excellent agreement with Schelkunoff's values, and the resonant length $0.936 \times \frac{1}{2}\lambda$ is only slightly less than the theoretical value $0.942 \times \frac{1}{2}\lambda$. Fig. 3 indicates that the input resistance depends on the centre spacing, increasing as the centre spacing is increased. The value of 61 ohms for the resonant resistance when the centre gap was reduced to $\frac{1}{2}$ in. is, however, in good agreement with Schelkunoff's value of 61·6 ohms.

* See p. 4.

The input impedance of a dipole exactly $\frac{1}{2}\lambda$ in length, $81 + j.50$ ohms, is high both as regards resistance and reactance compared with Schelkunoff's value $76 + j.44$ ohms. Further experiments, reported by these authors on the impedance of a dipole fed by bonded concentric lines, indicate the effect of the impedance of the bond on the measured impedance (fig. 4). This case is not directly comparable with the theory as we have given it since it corresponds to an aerial fed by two generators at finite distances on each side of its centre.

The experiments of Essen and Olliver were made with a wavelength of 60 cm., using a balanced line and standing wave indicator for impedance measurement. Various ratios of diameter to wavelength were used, and experimental curves are given showing the variation of impedance as the length of the aerial is varied from quite small values to 2λ. Their measurements near the resonant point $l \simeq \frac{1}{4}\lambda$ are not sufficiently accurate to give a good comparison with theory, but it is clear that their resonant resistances are very high, being all well over 70 ohms. This is undoubtedly because they used a comparatively large centre spacing. Their values for the input resistance at the anti-resonant point near $l = \frac{1}{2}\lambda$ are, however, in fairly good agreement with Schelkunoff's values as are their values for the length for zero reactance, within the accuracy of the experiments.

A certain amount of experimental work on the resonant lengths of full-wave and half-wave dipoles is quoted by Schelkunoff, but it must be admitted that much of the older experimental work is not sufficiently accurate to differentiate between the various theories. He also quotes some measurements on the input resistance of full-wave dipoles which seem to be in good agreement with his theoretical values. This is shown in fig. 22. All these measurements were made with unipoles above an earth, and the measured values have been doubled to correspond to the free-space dipole. Recent work has, however, shown how much care must be taken in the interpretation of the experimental work from the point of view of feeding conditions. Many feeding arrangements introduce lumped impedances at the feed point.

The most accurate recent measurements on unipoles appear to be those of Rösseler, Vilbig and Vogt(15), Brown and Woodward(14) and D. D. King(12, 13). The first authors used a comparatively long wave-length of about 36 m., the unipoles being supported on insulators above the earth by means of insulating guys. Impedance

measurement was by classical resonance methods. Their results mainly apply to quarter-wave unipoles and are in surprisingly good agreement with Schelkunoff's values even as regards input resistance of the resonant quarter-wave unipole, and impedance of a unipole exactly a quarter wave-length long. They studied the effect of coning the input end to the aerial in order to reduce the base capacity. The effect is small but significant, and probably may be accounted for by the change in shunt capacity between the base and earth. We shall return later to their measurements on fat unipoles.

The work of Brown and Woodward was carried out at a frequency of the order of 60 Mc./sec. The unipole was mounted above a metal earth plate and was fed from beneath the plate by a concentric feeder whose outer conductor was attached to the earth plate. For the fatter aerials investigated the overlap between the base and the earth plate was considerable and undoubtedly introduced a shunt capacity. The resonant lengths given for both half-wave and full-wave dipoles are considerably lower than the theoretical values and the rate of decrease with l/a much greater. This may be due to the feeding arrangements, but there seems to be an indication from a number of experiments that the theoretical value for the resonant length of a half-wave dipole is too large. The input resistance of a dipole $\frac{1}{2}\lambda$ long is in good agreement with theory, but a high value (of the order of 60 ohms) is found for the reactance. The values for the input resistance of full-wave aerials agree very well with Schelkunoff's values.

The effect of this type of feeding arrangement has been studied carefully by D. D. King[13] and discussed by Middleton and R. King[19]. The effect of the shunt capacity due to the gap is clearly demonstrated, and the care which is required in comparing experimental measurements with theoretical values is well brought out. The latest work by D. D. King seems to indicate better agreement with the Hallén type of approximation, as modified by Middleton and King, than with Schelkunoff's treatment. In particular, higher values for the resonant resistance of a half-wave dipole with $\lambda/4a$ about 100, of the order of 70 ohms, are found. Enough has already been said, however, to indicate that it would be unwise to assume that this is in itself a vindication of the theory, since this characteristic seems to be one most affected by input conditions. Indeed, the experiments of Smith and Holt Smith and Cochrane support the lower values of Schelkunoff.

Summarizing, one may say that in view of the difficulty in realizing the ideal conditions assumed by theory the agreement between theory and experiment is remarkably good, and the difference between Schelkunoff's values and those obtained from the later treatments of Hallén's equation are not very significant in practice. These theoretical values are certainly an excellent guide for a first attempt at a design and will probably only require slight modification to suit the particular feeding arrangement used. For example, base capacity may be approximately allowed for by electro-static considerations. In any case the final adjustments due to the presence of insulators, etc., are likely to be even larger than the discrepancies between the theoretical and actual values for the free aerial.

4·9. Mutual impedances

We now return to the calculation of mutual impedances. Although the simple sinusoidal theory fails to give the correct value for the electric field at the surface of a single aerial due to the current flowing in it, it gives of course the field at a distant point to a very high degree of approximation. In calculating the mutual impedance between two aerials 1 and 2 we require to know the field at the surface of 2 due to the current in 1. This is given with sufficient accuracy in most cases by the simple sinusoidal theory. The total electric field at the surface of 2 must, of course, vanish, but this is made up of a contribution from 2 itself and that due to 1. Let $E_{12}(x)$ be the component of the electric field along the aerial at a point at distance x from the centre of aerial 2, due to a current $i_1(x)$ in 1 having a value i_0 at the centre. Then by equation (42) of Chapter 3 the voltage e_2 developed at the centre of 2 due to the field $E_{12}(x)$ is given by

$$e_2 = \frac{1}{i_0} \int E_{12}(x)\, i_{2i}(x)\, dx, \tag{81}$$

where $i_{2i}(x)$ is the current which would be developed in 2 when transmitting, having the value i_0 at the centre. We may, to the usual degree of accuracy, replace $i_{2i}(x)$ by a sinusoidal current $i_s(x)$ having the value i_0 at the centre. Hence, assuming sinusoidal currents in both 1 and 2, we have

$$\frac{e_2}{i_0} = Z_{12} = -\frac{1}{i_0^2} \int E_{12}(x)\, i_s(x)\, dx. \tag{82}$$

Equation (79) gives good values for the mutual impedance [correct to order of $1/(2\log_e 2l/a)$] between two aerials whose length

is not far removed from $\frac{1}{2}\lambda$. Care must be taken when this does not apply, since the induced currents are of the same order of magnitude as the deviation of the main currents from a sinusoidal distribution. Also, as we shall see, a slightly different procedure is necessary when the length of the aerial is nearly equal to λ.

As an example we shall work out the mutual impedance of two half-wave dipoles placed end to end, which we have used to calculate the impedance of a full-wave aerial. The field E_{12} due to the current in aerial 1 at a point on its axis at distance x from the centre is obtained from (25) in the form

$$E_{12} = -j \cdot 30i_0 \left[\frac{e^{-jk(\frac{3}{4}\lambda + x)}}{\frac{3}{4}\lambda + x} + \frac{e^{-jk(\frac{1}{4}\lambda + x)}}{\frac{1}{4}\lambda + x} \right]. \tag{83}$$

Then

$$Z_{12} = j \cdot 30 \int_{-\frac{1}{4}\lambda}^{\frac{1}{4}\lambda} \cos kx \left[\frac{e^{-jk(\frac{3}{4}\lambda + x)}}{\frac{3}{4}\lambda + x} + \frac{e^{-jk(\frac{1}{4}\lambda + x)}}{\frac{1}{4}\lambda + x} \right] dx$$

$$= -15 \int_0^{4\pi} \frac{1 - e^{-jy}}{y} dy + 30 \int_0^{2\pi} \frac{1 - e^{-jy}}{y} dy \tag{84}$$

$$= 15[\gamma + \log_e \pi + \text{Ci}(4\pi) - \text{Ci}(2\pi)] + 15j[2\text{Si}(2\pi) - \text{Si}(4\pi)] \tag{85}$$

$$= \underline{25 \cdot 1 + j \cdot 19 \cdot 8} \text{ ohms.} \tag{86}$$

Numerical values for the mutual impedance between two half-wave dipoles are given in tables 3 and 4 of Chapter 2. Analytical expressions have been given by Carter[1] and various curves by Brown[2]. The latter, however, contain some errors, and the curve for a full-wave aerial is quite wrong. The formulae given by Carter may be simplified to some extent as the separation of the aerials increases beyond about 7λ (as has been shown by Appanasiev[20]). When the dipoles are parallel and their centres lie on the perpendicular bisector of each, the mutual impedance has a particularly simple form. When the separation d is greater than about 5λ the component of E along the dipole is nearly constant and is given by equations (8) and (10) of Chapter 3. We have, putting $\theta = 0$,

$$Z_{12} = \frac{60}{d} \exp\left[-j\left(\frac{2\pi d}{\lambda} - \frac{\pi}{2}\right) \right] \int_{-\frac{1}{4}\lambda}^{\frac{1}{4}\lambda} \cos kx \, dx$$

$$= \frac{60\lambda}{\pi d} \exp\left[i\left(\frac{\pi}{2} - \frac{2\pi d}{\lambda}\right) \right]. \tag{87}$$

The first-order corrections to the mutual impedance corresponding to Hallén's first-order approximation have been given by

King and Harrison(21), who have worked out in detail the case of two exactly similar aerials, placed symmetrically with regard to one another, one acting as a driver and the other as a parasite. For a value of $2 \log_e 2l/a$ of the order of 10 the first-order corrections amount to 2 or 3 ohms. Apart, however, from working out the exact form of polar diagrams near interference minima, the mutual impedances obtained from the sinusoidal approximation are normally sufficiently accurate.

4·9·1. *Mutual interaction of full-wave dipoles*

For full-wave dipoles equation (85) becomes invalid to the approximation given by the sinusoidal theory as the current at the centre of the aerial becomes zero. We must therefore proceed in
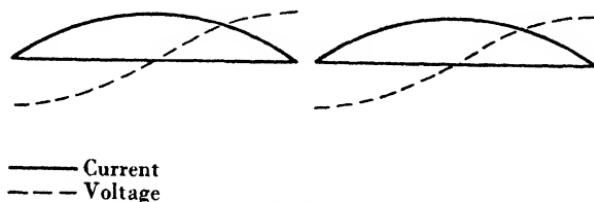
———— Current
— — — Voltage

Fig. 26. Current and voltage in full-wave driven aerial.

a slightly different manner, similar to that used to obtain the impedance of a full-wave dipole. In the case of full-wave dipoles it is more convenient to work with admittances. Linear equations defining the self- and mutual admittances may be written down in the same way as for the impedances. They are

$$i_1 = Y_{11}e_1 + Y_{12}e_2, \quad i_2 = Y_{12}e_1 + Y_{22}e_2, \tag{88}$$

where e_1, i_1 and e_2, i_2 are the currents and voltages at the inputs of the aerials 1 and 2. The current and voltage in a full-wave aerial are in the form shown in fig. 26, if the aerial is driven. Let i_{01} be the value of the current maximum in aerial 1 and i_{02} that in aerial 2. Then to the approximation of the sinusoidal theory using constant characteristic impedance Z_0 we have

$$e_1 = Z_0 i_{01}, \quad e_2 = Z_0 i_{02}. \tag{89}$$

Let $i_1(x)$ be the form of the current in aerial 1 and $i_2(x)$ that in aerial 2. Let $i_1(x)$ produce an electric field having component $E_{12}(x)$

along 2 and, similarly, $i_2(x)$ will produce $E_{21}(x)$ along 1. Then we have, by considering the 'mutual' part of the complex power,

$$Y_{12}e_1e_2 = -\int_{\text{along } 2} E_{12}(x)\,i_2(x)\,dx. \tag{90}$$

Using (86) we have

$$Y_{12}Z_0^2 = \frac{1}{i_{01}i_{02}}\int_{\text{along } 2} E_{12}(x)\,i_2(x)\,dx. \tag{91}$$

Values of $Y_{12}Z_0^2$ have been calculated by Scott* for two full-wave dipoles and are shown in table 5 of Chapter 2.

——— Symmetrical current distribution
— — — Anti-symmetrical current distribution
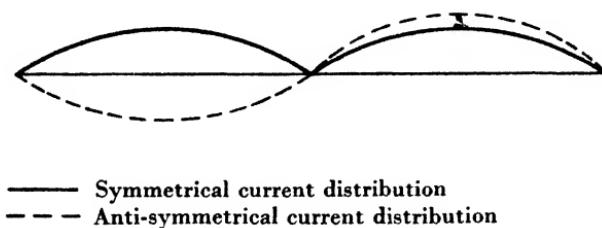
Fig. 27. Current in full-wave dipole.

Care must be taken, however, in using these. In particular, it is essential that the aerial arrangement is such that the symmetrical type of current shown in fig. 26 is produced. As an example of how one may be led into error consider the case of a driven full-wave dipole with a parallel full-wave dipole shorted at the centre. The shorted dipole can resonate in a mode carrying an anti-symmetrical current as shown in fig. 27. Strictly, according to the above calculation the current in the parasite should be zero, but this applies to the symmetrical form of the current only. The shorted full-wave dipole is not a true resonant parasite, and the analogue of the half-wave parasite is the open-circuited full-wave dipole which will resonate with symmetrical current form.

REFERENCES

(1) Carter, P. S. Circuit relations in radiating systems and applications to antenna problems. *Proc. Inst. Radio Engrs, N.Y.*, **20**, 1004 (1932).

(2) Brown, G. H. Directional antennae. *Proc. Inst. Radio Engrs, N.Y.*, **25**, 78 (1937).

* Unpublished calculations by J. M. C. Scott, performed for Ministry of Supply.

(3) King, L. V. On the radiation field of a perfectly conducting base-insulated cylindrical antenna, and the calculations of radiation resistance and reactance. *Philos. Trans.* A, **236**, 381 (1937).

(4) Schelkunoff, S. A. Theory of antennas of arbitrary size and shape. *Proc. Inst. Radio Engrs, N.Y.*, **30**, 493 (1942).
Schelkunoff, S. A. Principal and complementary waves in antennas. *Proc. Inst. Radio Engrs, N.Y.*, **34**, 23 P (1946).

(5) Burgess, R. E. The relationship between the transmitting and receiving characteristics of aerials. *Wireless Engr*, **21**, 154 (1944).

(6) Hallén, E. Theoretical investigations into the transmitting and receiving qualities of antennas. *Nova Acta Uppsala*, ser. IV, **11**, no. 4 (1938).

(7) King, R. and Blake, F. G. The self-impedance of a symmetrical antenna. *Proc. Inst. Radio Engrs, N.Y.*, **30**, 335 (1942).

(8) Abraham, M. Die elektrischen Schwingungen um einen Stabförmigen Leiter, behandelt nach der Maxwellschen Theorie. *Ann. Phys., Lpz.*, **66**, 435 (1898); **2, 3** (1900).

(9) Page, L. and Adams, N. L. The electrical oscillations of a prolate spheroid. *Phys. Rev.* **53**, 819 (1938).

(10) Chu, L. J. and Stratton, J. A. Forced oscillations of a prolate spheroid. *J. Appl. Phys.* **12**, 241 (1941).

(11) Chu, L. J. and Stratton, J. A. Forced oscillations of a conducting sphere. *J. Appl. Phys.* **12**, 236 (1941).

(12) King, D. D. Microwave impedance measurements with application to antennas. *J. Appl. Phys.* **16**, 445 (1945).

(13) King, D. D. Dipole antennas. *J. Appl. Phys.* **12**, 844 (1946).

(14) Brown, G. H. and Woodward, O. M. Experimentally determined characteristics of cylindrical antennas. *Proc. Inst. Radio Engrs, N.Y.*, **33**, 257 (1945).

(15) Rösseler, G., Vilbig, F. and Vogt, K. Über das elektrische Verhalten von Vertikalantennen in Abhängigkeit von ihrem Durchmesser. *T.F.T.* **28**, 170 (1939).

(16) Schelkunoff, S. A. Concerning Hallén's integral equation for cylindrical antennas. *Proc. Inst. Radio Engrs, N.Y.*, **33**, 872 (1945).

(17) Gray, Marion C. A modification of Hallén's solution of the antenna problem. *J. Appl. Phys.* **15**, 61 (1944).

(18) Bowkamp, C. J. Hallén's theory for a straight perfectly conducting wire, used as a transmitting or receiving aerial. *Physica*, **9**, 609 (1942).

(19) Middleton, D. and King, R. The thin cylindrical antenna—a comparison of theories. *J. Appl. Phys.* **17**, 273 (1946).

(20) Appanasiev, K. J. Simplifications in the consideration of mutual effects between half-wave dipoles. *Proc. Inst. Radio Engrs, N.Y.*, **34**, 635 (1946).

(21) King, R. and Harrison, C. W. Mutual and self-impedance for coupled antennas. *J. Appl. Phys.* **15**, 481 (1944).

Chapter 5

REFLECTORS AND DIRECTORS

5·0. Uses of reflectors and directors

In many aerial problems it is required to provide radiation only in 180° of azimuth which we shall call the 'forward direction' and to screen from radiation the remaining 180°, in the 'backward direction'. For this purpose reflectors are used. For very short wave aerials these frequently consist of metallic screens, overlapping the aerial to some extent to cut down diffracted radiation. On longer wave-lengths, however, these screens become unwieldy and parasitic reflectors are normally used. These consist of continuous dipoles which are nearly resonant, having a length just greater than $\frac{1}{2}\lambda$, spaced at a distance usually less than a quarter of a wave-length behind the driven elements. Such undriven elements are called 'parasitic' elements.

5·1. Single driver and parasite

We shall consider first a single driver and parasite as shown in fig. 28. Suppose the parasitic element is placed at a distance $\frac{1}{4}\lambda$ behind a driven element, and its self-reactance is adjusted by altering its length so that the current induced in it leads the current in the driven element by 90°, then the radiation from the two elements at right angles to them, looking towards the parasite, will be 180° out of phase with each other. This is because of the extra 90° lead of the radiation from the parasite arising from the $\frac{1}{4}\lambda$ spacing. If the currents in parasitic and driven elements are nearly equal, there will be approximate, but not quite complete, cancellation of radiation in the backwards direction. In the opposite direction the radiation from the two elements will be in phase and will add. This indicates another use of reflectors, to enhance the radiation in the 'forward direction'. Unfortunately, as we shall see, the optimum conditions for diminution of backward radiation are not those for maximum increase of forward radiation. It is found that when parasitic elements are slightly longer than $\frac{1}{2}\lambda$ the current in them leads the current in the driver and they act as reflectors. When somewhat less than $\frac{1}{2}\lambda$ the current lags behind

that in the driver and radiation is enhanced in the direction looking towards the parasite. In this case the element is called a 'director'.

The equations for the current in a driver and single parasite have already been given in Chapter 2. The current i_2 in the parasite is given in terms of the current i_1 in the driver by

$$i_2 = -Z_{12}i_1/Z_{22}, \tag{1}$$

Z_{12} being the mutual impedance and Z_{22} the self-impedance of the parasite. Let $Z_{12} = r_1 + jx_1$ and $Z_{22} = r_0 + jx$. As the length of the parasite is slightly varied, as we have seen in Chapter 4, only x will vary appreciably, the rate of change of x with length being given by equation (34) of Chapter 4. In order to see what happens when x is varied let us take the simplified case when Z_{12} is purely resistive (for example, spacing $d = 0.12\lambda$, see table 3).

Then

$$-\frac{r_1 i_1}{r_0 + jx} = \frac{r_1}{(r_0^2 + x^2)^{\frac{1}{2}}} e^{j(\pi - \gamma)} i_1,$$

where $\tan \gamma = x/r_0$. When $x > 0$, i_2 leads i_1 by an angle lying between π and $\frac{1}{2}\pi$. When $x < 0$, i_2 leads i_1 by an angle lying between π and $\frac{3}{2}\pi$, or lags i_1 by an angle lying between π and $\frac{1}{2}\pi$. Since $(r_0^2 + x^2) > r_1^2$, $i_2 < i_1$, and complete cancellation does not take place. The best value for x to make the parasite a reflector will be such as to give approximately

$$(\pi - \gamma) + \frac{2\pi d}{\lambda} = \pi, \tag{2}$$

or $\gamma = 43°$, giving $x = 68$ ohms if we take $r_0 = 73$ ohms. The optimum condition for a director will be approximately given by

$$(\pi - \gamma) - \frac{2\pi d}{\lambda} = 0 \quad \text{or} \quad \gamma = 137°.$$

This gives $x = -78$ ohms.

These values are only approximate and are meant to illustrate the type of behaviour we may expect. The general expressions for the radiation from the two elements are somewhat complex, and their behaviour is best studied from graphs.

Let θ be the angle between a given direction in the plane of the elements and the line joining their centres. We shall assume that the elements are parallel and symmetrically placed with regard to each other (fig. 28). We shall take $\theta = 0$ as the direction looking

towards the driver from the parasite along the centre line. The combined field strength from the elements is in the form

$$E = \text{const.} f(\theta)(i_1 + i_2 e^{-jkd\cos\theta}), \tag{3}$$

where $f(\theta)$ is the polar diagram of each element obtained from equation (10), Chapter 3 and $k = 2\pi/\lambda$. The polar diagram function $F(\theta)$ is given by

$$F(\theta) = \text{const.} \times f(\theta) \left| i_1 + i_2 e^{-jkd\cos\theta} \right|. \tag{4}$$

The front-to-back ratio is given by

$$\left| \frac{i_1 + i_2 e^{-jkd}}{i_1 + i_2 e^{jkd}} \right|, \tag{5}$$

i_2/i_1 being given by (1).
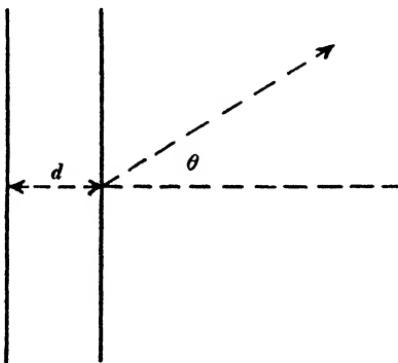
Fig. 28. Driven element and parasite.

By calculating the input impedance from equations (1) of Chapter 2, using (1), the actual power-gain may be calculated for the forward and backward directions, as shown in §2·3. This quantity is frequently more important than the simple ratio (5). Curves showing the power-gain and polar diagram for a large number of values of d as x is varied have been given in a paper by Walkinshaw(1) on Yagi aerials. Similar curves have been given for specific examples in the literature, for example by Brown (Chapter 4, Ref. (2)). Some typical curves are given in fig. 29, showing power-gain and input resistance as x is varied for a number of values of d. Polar diagrams for a great variety of values of x and d have been given by Walkinshaw. In fig. 30 we show typical examples of polar diagrams for $d = 0·15\lambda$. The optimum power-gain is of the order of 3

when $x = 20$ ohms and when the parasite is acting as a reflector with a spacing of about 0.15λ. Fig. 29 indicates that a greater power-gain might be expected by decreasing d below the value of 0.1λ. In fact,
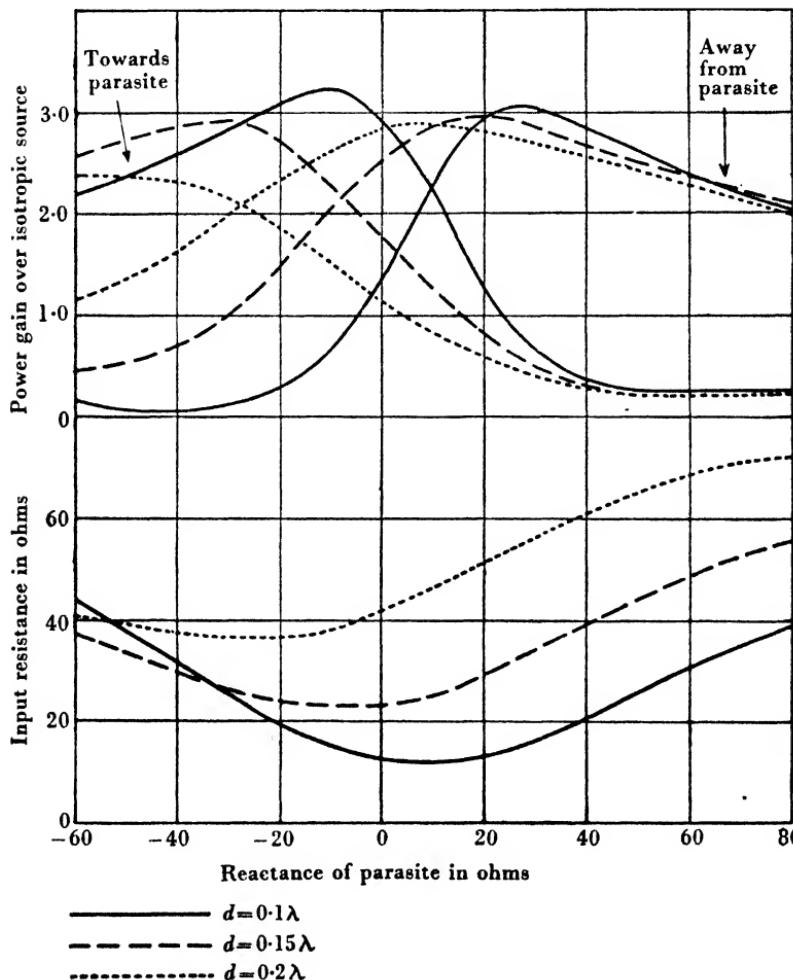
Fig. 29. Power-gain and input resistance of aerial consisting of driver and single parasite acting as director or reflector.

the best value is about 0.1λ, which gives a power-gain of about 3.1. As the spacing d becomes of the order of 0.1λ and less the input resistance becomes inconveniently low, as will be seen from fig. 29. For a spacing of 0.15λ the best value of x for a director is about

-35 ohms, giving a gain of 2·8. Again, the gain may be increased to about 3·25 by decreasing d to 0·1λ. For a reflector to give the best compromise between reduction of backward radiation and forward gain $x = 45$ ohms. The forward gain is then about 2·5, and the input resistance about 45 ohms. The front-to-back ratio is 3·6, and this is about the smallest value practically available. It may be
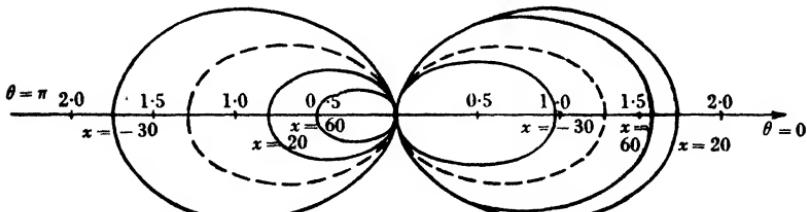
(a) Plane of elements
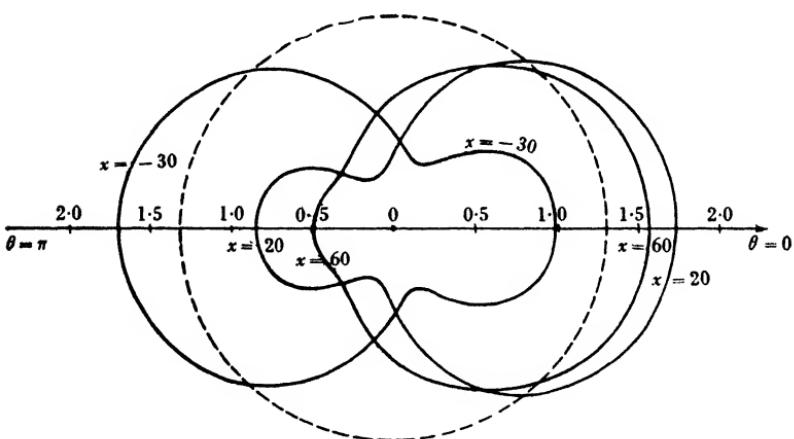
(b) Plane perpendicular to elements

Fig. 30. Polar diagrams for single driver and parasite acting as reflector, and director. Field strength is relative to isotropic source of same power, spacing is 0·15λ. Dotted curve is for $\lambda/2$ dipole without parasite.

noted that the input reactance may be removed by variation of the length of the driven element. For a reflector the backward radiation consists of a single lobe with a maximum on the centre line and will be a minimum if the radiation in this direction is reduced to a minimum. This is not true for more complex arrangements.

The driver and single parasite have been studied experimentally by McPetrie and Saxton(2). The general behaviour predicted by theory has been well verified. (Note that these writers give their

power-gain relative to a half-wave dipole.) The optimum power-gain found for a director was 4·3, with a spacing of 0·1λ. This is somewhat higher than predicted by the theory as we have given it. Their polar diagram for the reflector agrees well with that shown in fig. 30, but that found for a director shows some curious lobes in the backward direction, not predicted by theory. The unsymmetrical character of their diagram, however, appears to indicate some source of interference.

5·2. Two driven elements with parasites

We shall next consider the use of parasitic reflectors with an aerial array, taking as the simplest case a vertical stack consisting of two half-wave dipoles separated by $\frac{1}{2}\lambda$ and fed with equal currents,
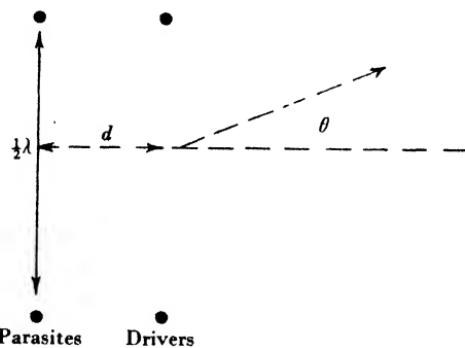
Fig. 31. Aerial array consisting of two driven elements and two parasites.

having equal parasitic elements placed symmetrically at the same level and at a horizontal distance d behind them. The arrangement is shown in fig. 31. The linear equations representing this system contain the four unknown currents in the elements, it being assumed that the driving voltage is given. By symmetry these may, however, be reduced to two. After these equations have been solved for the currents the power-gain, input impedance and the polar diagram may be calculated. The expressions for these are somewhat complex and will not be given here(1). Typical curves of power-gain and input resistance are, however, shown in fig. 32. In fig. 33 are shown polar diagrams of the backward radiation when the parasites act as efficient reflectors. By comparing figs. 29, 32 and 30, 33 it will be seen that two important differences exist between the two-tier and
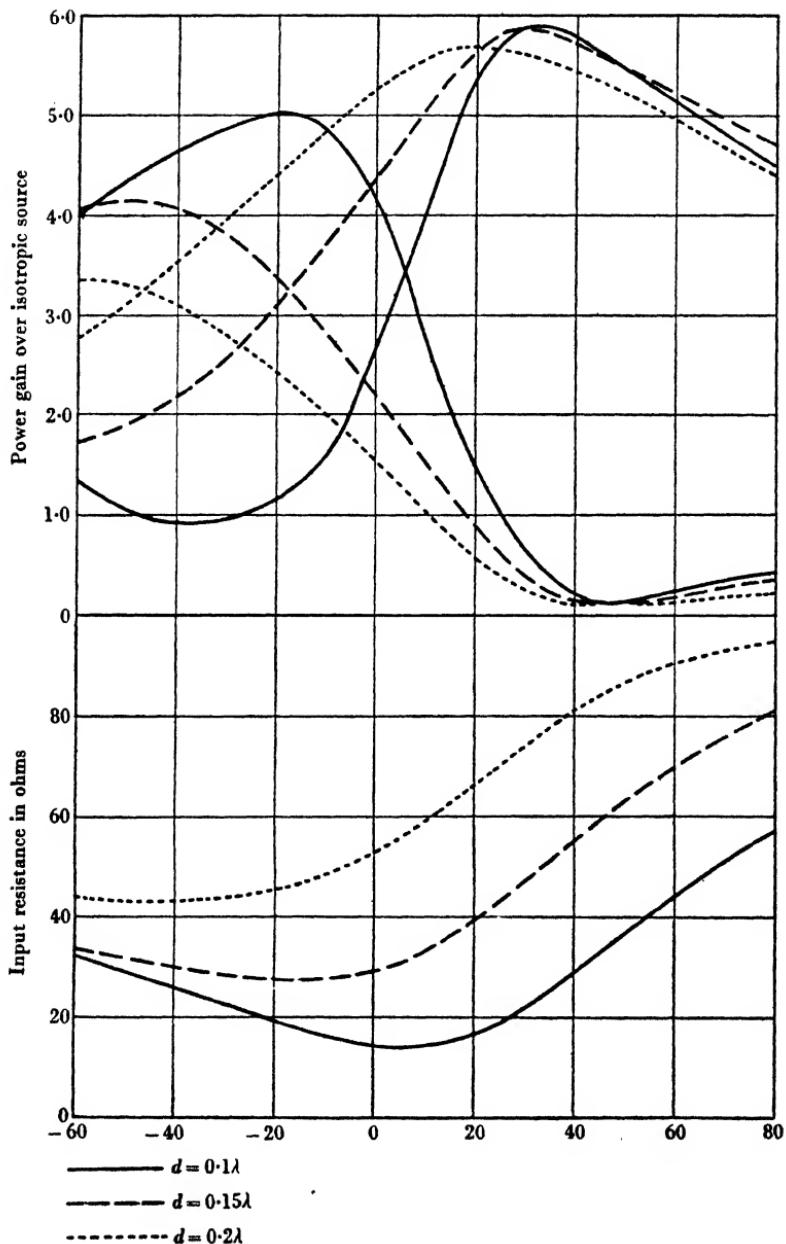
Fig. 32. Power-gain and input resistance of two-tier array
with parasitic elements.

(a) Plane of elements

(b) Plane perpendicular to elements
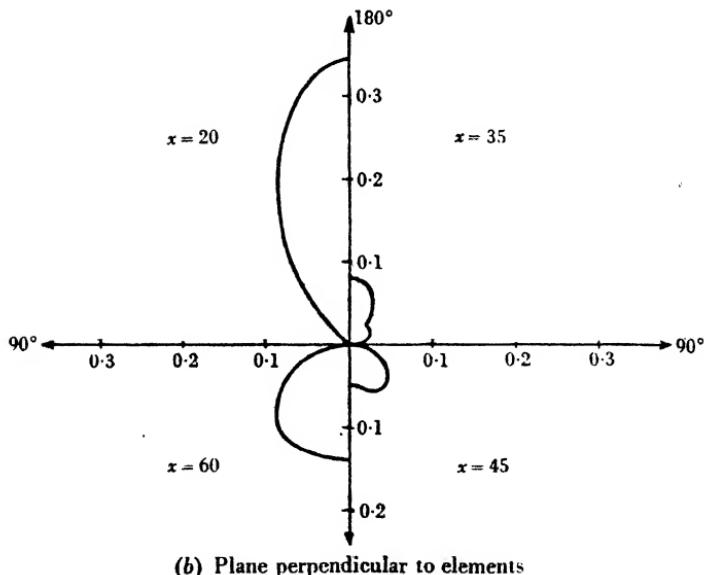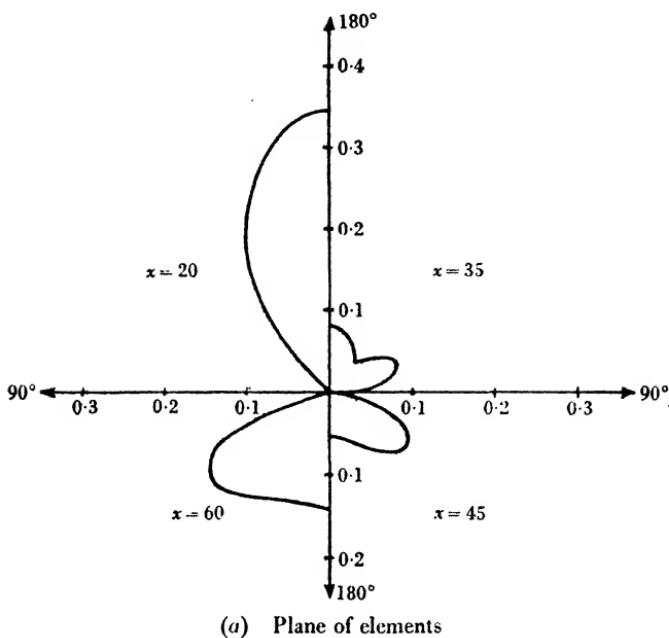
Fig. 33. Polar diagrams of backward radiation of two-tier array with parasitic reflectors. (Field strength is relative to maximum of forward radiation.)

single-tier arrays. First the available front-to-back ratios are much larger, i.e. the reflectors are acting much more efficiently. For a spacing of 0.15λ with a reactance of about 45 ohms a front-to-back ratio of about 20 to 1 may be obtained. This, however, does not represent the true efficiency of the reflectors as we shall see. The forward gain with this condition is about 5·75, the maximum available being just less than 6·0. The largest gain available with directors at this spacing is 4·25 and occurs when $x = -45$ ohms. For a 0.1λ spacing a gain of 5 is available with $x = -20$ ohms. It will be noted that for the two-tier array more gain is available by the use of the parasitic elements as reflectors than as directors. The reverse is true for the single-tier array.

The second marked difference between the two- and one-tier arrays may be noted in the polar diagrams of the backward radiation when the parasites are acting as efficient reflectors. The backward radiation ceases to have its maximum on the centre line and splits up into two or three lobes. The condition for minimum radiation in the backward direction along the centre line exists when a minimum between two lobes lies on this line. However, this is not the best condition for general reduction of backward radiation which exists when we have three lobes of equal size—the so-called 'clover-leaf' pattern. This occurs when $x = 35$ ohms.

These considerations indicate that a new definition of 'front-to-back' ratio is required. The most reasonable one would appear to be 'ratio of field strength in the direction of the maximum of the forward radiation to that in the direction of the maximum of the backward radiation'. The ratio of field strengths at the front and back maxima for the optimum condition will be seen to be about 12 to 1. This should be compared with the value 20 to 1 obtained by considering the centre line only. This behaviour has been verified by W. Cochrane* in considerable detail. The patterns shown in fig. 33 were all found in the correct sequence, the only difference being that the minimum found in the polar pattern on the centre line in the backward direction when the reflectors were adjusted for minimum radiation in this direction was even deeper than predicted by theory. This is precisely where small errors in the mutual impedances would be important. Thus we see that measurements of front-to-back ratio made on the centre line only give

* Unpublished wartime research for the Ministry of Aircraft Production.

a completely false indication of the amount of radiation in other backward directions.

The effect of insulators supporting the parasite has been examined by various writers(3, 4). The insulators have the effect of reducing the length of the parasite required for optimum operation as a reflector. It has been shown(3) that when the effect of a particular kind of insulator is known at one frequency it can be deduced for a fairly wide range of frequencies by assuming that the insulator is equivalent to a fixed length of the aerial wire.

5·3. Multi-element arrays

Calculations have also been made for arrays having four and six tiers, and the results differ surprisingly little from those found for the two-tier array.* The same form of backward polar diagram is found. These calculations were carried out with 0·17 spacing. For a six-tier array the 'clover-leaf' pattern occurs with $x = 35$ ohms, and the optimum front-to-back ratio as defined above is of the order of 12 to 1. The power-gain for a four-tier array under these conditions is 15 and for a six-tier array is 26. The relation between power-gain and spacing is not critical, and table 6 shows the approximate variation of power-gain with n of an n-tier array of drivers spaced $\frac{1}{2}\lambda$ apart with parasitic reflectors behind at a distance of the order of $0·15\lambda$.

Table 6. *Power-gain of n-tier array with reflectors*

No. of tiers	1	2	3	4	5	6
Power-gain over isotropic source	3	6	10	15	20	26

5·4. Variation of reflector efficiency with frequency

For arrays having more than one tier the efficiency of reflectors begins to fall off seriously when x departs by more than about 15 ohms from the optimum value. From fig. 24 it will be seen that for half-wave elements this would only correspond to about a 1·5 % change in frequency when $\lambda/a = 1200$ and 0·75 % change when $\lambda/a = 40,000$. This shows up one of the main disadvantages of parasitic or 'tuned' reflectors and raises the need for reflectors which are relatively aperiodic. The 'band-width' properties of tuned

---

* Unpublished calculations by H. G. Booker and W. Walkinshaw, made for the Ministry of Aircraft Production.

reflectors may of course be improved by using fat parasitic dipoles. A more effective method is to use more than one parasite associated with each driven element as in the so-called 'trigonal reflector'. Two such arrangements which have been used are shown in fig. 34 (a) and (b) by means of which a considerable improvement over the single-element reflectors may be obtained. Nevertheless, it is clear that whenever practically possible some form of aperiodic reflector has distinct advantages. For wave-lengths greater than a few metres these are apt to be clumsy, but for shorter wave-lengths are frequently more convenient.
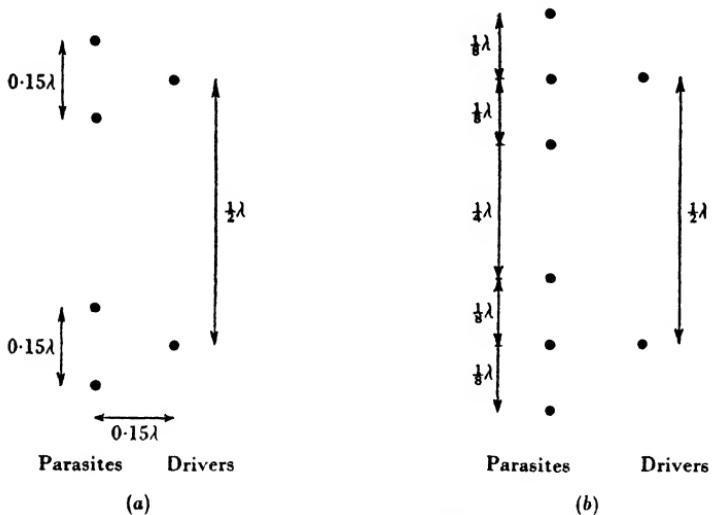
Fig. 34. Multi-element reflectors.

5·5. Aperiodic reflectors

An aperiodic reflector simply consists of a conducting screen placed behind the aerial. The best screen consists, of course, of a sheet of material of high conductivity and theoretically should be infinite in extent if it is to cut off all radiation in a backward direction. If this is not so radiation will leak round by diffraction. Calculations by N. F. Mott* using Sommerfeld's diffraction formula(5) have shown that, provided the spacing between the sheet and aerial is of the order of $\frac{1}{8}\lambda$, an overlap of $\frac{1}{2}\lambda$ all round will be sufficient to provide a very effective screen. This overlap is, of course, most

* Unpublished calculations for Ministry of Supply.

essential in the direction at right angles to the dipoles, a smaller overlap at the 'end-on' position being tolerable. With such an arrangement front-to-back ratios of the order of 50 to 1 are obtainable.

5·5·1. *Diffraction at a straight edge*

Mott's calculations consisted in evaluating the field due to a dipole placed at distance ρ from the edge of a semi-infinite screen and at perpendicular distance p from the screen. When $p \ll \rho$ the formula for the radiation in a backward direction at right angles to the screen reduces to a fairly simple form. When the dipole is parallel to the screen and parallel to the edge the ratio of forward to backward radiation at right angles to the screen is

$$2\pi 2^{\frac{1}{2}}\rho^{\frac{1}{2}} \sin (2\pi p/\lambda)/p \, \lambda^{\frac{1}{2}}. \tag{6}$$

For small values of p this tends to a constant value and there is no advantage to be gained in reducing p below $\frac{1}{8}\lambda$, a value frequently used in practice as giving a good compromise between power-gain and input resistance. For $p = \frac{1}{8}\lambda$ and $\rho = \frac{1}{2}\lambda$, (6) has the value 22. When the dipole is perpendicular to the edge and parallel to the screen the backward radiation has half the value given by (6) when $p \ll \rho$.

5·5·2. *Mesh and wire screens*

For many purposes the high wind resistance of a sheet of metal renders it inconvenient as a screen, and a wire mesh or set of parallel wires may be used instead. Provided the wires are parallel to the electric vector and are spaced at a distance d apart, much smaller than a wave-length, they act as an efficient screen. An approximate formula for the transmission coefficient of an infinite screen of such wires of radius a for plane waves incident normally has been given by N. F. Mott:[*]

$$\gamma = (2d/\lambda) \log_e (d/2\pi a). \tag{7}$$

The problem of the diffraction of plane waves by such an infinite screen has been treated more exactly by G. G. Macfarlane[6], who obtains a more general formula for the case of oblique incidence. Practical experience with parallel wire screens has shown that the value given by (7) is substantially correct.

* Unpublished calculations for Ministry of Supply.

Screens made up of parallel wires are not very convenient and 'chicken netting' is more frequently used. Although this is not amenable to exact treatment an approximate value for the transmission coefficient may be obtained by considering the screen as equivalent to one having horizontal wires spaced with the mesh dimensions as the wires in the mesh at right angles to the electric field will not contribute appreciably. Such screens have, as we shall see, been widely used for radar aerials with wave-lengths of the order of $1\frac{1}{2}$ m., and the above considerations both as regards transmission and diffraction have been well borne out in practice.

5·6. Effect of reflector screen on input impedance and power-gain

For the purpose of calculating the power-gain and input impedance of an aerial array backed by an efficient reflecting screen, the screen may be regarded as infinite and perfectly conducting. It may therefore be replaced by an aerial array which is the image of the array in front of the screen. If the elements of the original array are parallel to the screen the currents in the image are equal in magnitude but 180° out of phase. By use of mutual impedances the currents in the two arrays may then be calculated and hence the input impedance, power-gain and polar diagram.

For example, let us consider a single driven element placed parallel to the reflecting sheet at a distance d from it. Let i be the current at the mid-point of the element and e the voltage at the driving point. Then clearly, in the usual notation,

$$e = (Z_{11} - Z_{12}) i. \qquad (8)$$

The input impedance Z_i is given by

$$Z_i = Z_{11} - Z_{12}(2d). \qquad (9)$$

Z_i tends to zero as d tends to zero. The input power P is

$$R_e(\tfrac{1}{2}ei) = \tfrac{1}{2}(r_{11} - r_{12}) |i|^2.$$

The field strength E in the plane of the element and its image is given by

$$E = 2E_i f(\theta) \sin\left(\frac{2\pi d \cos\theta}{\lambda}\right), \qquad (10)$$

where E_i is the field strength in the equatorial plane of the element when it carries a current i, and $f(\theta)$ its polar diagram

$$\cos(\tfrac{1}{2}\pi \sin\theta)\cos\theta.$$

If d is so small that $\sin(2\pi d/\lambda)$ may be replaced by $2\pi d/\lambda$, and we use the approximate form $f(\theta) = \cos\theta$, the polar diagram function in the plane of the elements is simply $\cos^2\theta$. We note the narrowing due to the screen. In the plane at right angles to the elements the field strength is given by (10) with the factor $f(\theta)$ replaced by unity. For small d, the polar diagram function is then just $\cos\phi$. On the normal to the screen $\theta = 0$ and the field strength E_0 is given by

$$E_0 = 2E_i \sin\left(\frac{2\pi d}{\lambda}\right). \tag{11}$$

Now a single element carrying current i will have input power $\frac{1}{2} = r_{11}|i|^2$. Comparing the field strengths for the same input power we see that the power-gain G of the element with reflector is given by

$$G = \frac{4r_{11}}{[r_{11} - r_{12}(2d)]} \sin^2\left(\frac{2\pi d}{\lambda}\right) G_0, \tag{12}$$

where G_0 is the power-gain of a single element without reflector. We note that $G = 0$ when $d = \frac{1}{2}\lambda$, as would be expected, since the radiation from the element and its image would be in anti-phase in the direction of the normal. Since $r_{11} - r_{12}$ tends to zero like const. $\times d^2$ we note that G tends to a constant value as $d \to 0$. There is thus no advantage to be gained by making d very small. The variation of power-gain and input resistance with d is shown in fig. 35.

For an array consisting of two elements fed in phase and $\frac{1}{2}\lambda$ apart the input impedance Z_i is given by

$$Z_i = \frac{1}{2}[Z_{11} + Z_{12}(\frac{1}{2}\lambda) - Z_{12}(2d) - Z_{12}(f)], \tag{13}$$

where $f^2 = \frac{1}{4}\lambda^2 + 4d^2$. The power-gain G is given by

$$G = \frac{4[r_{11} + r_{12}(\frac{1}{2}\lambda)] \sin^2(2\pi d/\lambda)}{r_{11} + r_{12}(\frac{1}{2}\lambda) - r_{12}(2d) - r_{12}(f)} G_0, \tag{14}$$

where, as before, G_0 is the power-gain in the absence of the screen (see § 2·3). The input resistance and power-gain for this array are also shown in fig. 35.

When reduction of backward radiation is not a primary consideration a spacing d of about 0.725λ for a single element gives the best gain and also a reasonable input resistance of about 83 ohms.

More gain may be had for small values of d, but with a much lower value of input resistance which will make matching more critical and ohmic losses relatively more important. For a two-tier array
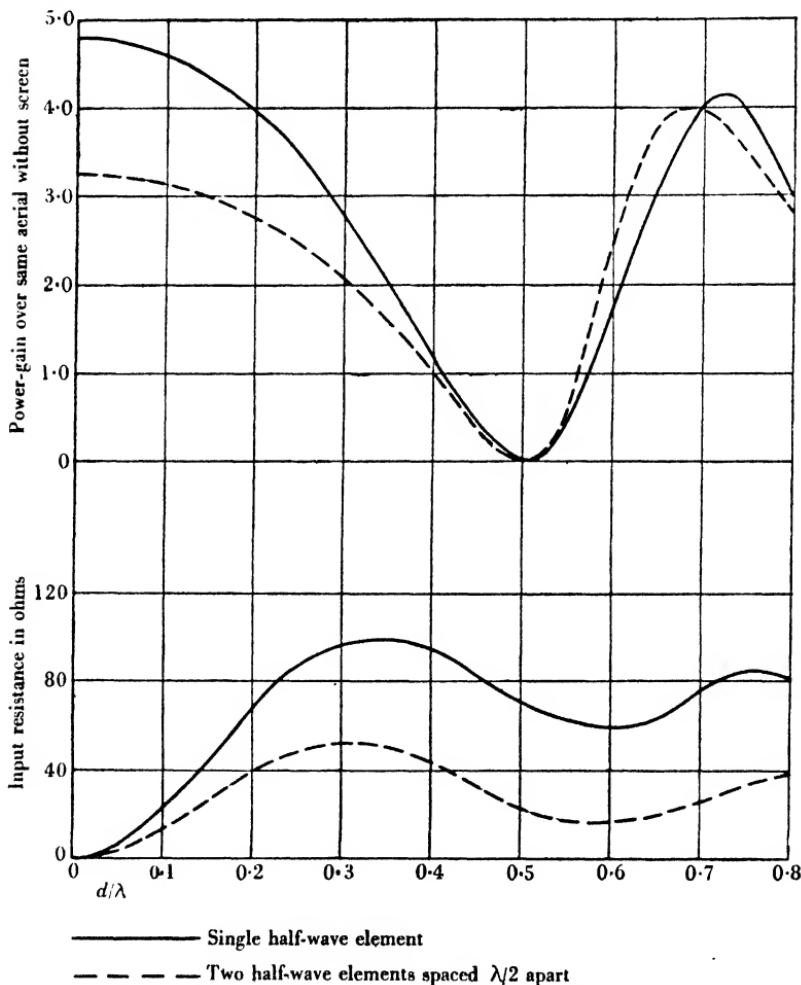
Fig. 35. Power-gain and input resistance of aerials backed by reflector screen.

the optimum spacing is about 0.65λ. Note that this gives more gain than we get with small values of d in contrast to what is found for a single-tier array.

The polar diagram function $F_1(\theta)$ in a plane parallel to that through

a driven element and its image is given for values of θ corresponding to forward radiation and not very nearly equal to 90° by

$$F_1(\theta) = \sin\left(\frac{2\pi d \cos\theta}{\lambda}\right) f(\theta) \bigg/ \sin\left(\frac{2\pi d}{\lambda}\right). \tag{15}$$

If d/λ is so small that we may write $\sin(2\pi d/\lambda)$ approximately equal to $2\pi d/\lambda$, then on substituting the approximate form $f(\theta) = \cos\theta$ we have

$$F_1(\theta) \simeq \cos^2\theta. \tag{16}$$

In the plane at right angles to the elements the polar diagram $F_2(\phi)$ is given by

$$F_2(\phi) = \cos\left(\frac{\pi \sin\phi}{2}\right)\sin\left(\frac{2\pi \cos\phi}{\lambda}\right)\bigg/\sin\left(\frac{2\pi d}{\lambda}\right), \tag{17}$$

where ϕ is the angle made with the normal to the screen. Again when $2\pi d/\lambda$ is small this may be written in the approximate form

$$F_2(\phi) \simeq \cos^3\phi. \tag{18}$$

It is of interest to note that the polar diagram in the latter plane is narrower than that in the former in contrast to what one finds for a single element. This is because of the factor $\cos\left(\dfrac{\pi \sin\phi}{2}\right)$ arising from the interference between elements $\frac{1}{2}\lambda$ apart, which tends more rapidly to zero as ϕ tends to $\frac{1}{2}\pi$ than does the function $f(\phi)$.

REFERENCES

(1) Walkinshaw, W. Theoretical treatment of short Yagi aerials. *J. Instn Elect. Engrs*, **93**, IIIA, 598 (1946).
(2) McPetrie, J. S. and Saxton, J. A. Some experiments with linear aerials. *Wireless Engr*, **23**, 107 (1946).
(3) Smith, R. A. and Holt Smith, C. The elimination of errors from crossed dipole direction finding systems. *J. Instn Elect. Engrs*, **93**, IIIA, 575 (1946).
(4) Saxton, J. A. and Ford, L. H. Dipole reflector insulation. *Wireless Engr*, **23**, 325 (1946).
(5) Baker, B. B. and Copson, E. T. *Huygens' Principle*. Oxford (1939).
(6) Macfarlane, G. G. Surface impedance of an infinite parallel wire grid at oblique angles of incidence. *J. Instn Elect. Engrs*, **93**, 1523 (1946).

Chapter 6

HIGH-POWER TRANSMITTING AERIALS FOR WAVE-LENGTHS OF A FEW METRES

6·0. Introduction

In the following chapters we shall deal with various applications of the principles which have been discussed. We shall be mainly concerned with matters of general importance from a design point of view rather than with the description in detail of particular aerial systems. Aerials are such that it is very difficult to standardize a design, and each application requires an aerial system specially adapted to meet its peculiar needs. There are, however, many principles of design which are common to all.

6·1. Radar transmitting aerials

Aerials for early warning radar were normally required to operate on a 'spot' frequency. The nature of the pulse transmission, however, required that the aerials should have a band width sufficient to accommodate the spectrum of the pulse. Five microseconds was about the shortest duration of pulse used in the metre band, so a band width of at least 500 kcyc./sec. was desirable[1]. The aerials used did not fall off appreciably in efficiency in this band which corresponds to $\pm 1 \%$ at 25 Mcyc./sec. Parasitic reflectors are the most sensitive elements in a broadside array, and this is just about the limit of their efficient operation if they are made from 200 lb./mile copper wire and hence have $\lambda/d = 3000$. 1% frequency change therefore corresponds to a change in reactance of 14 ohms (fig. 24), which as we have seen in Chapter 5 begins to make the reflector inefficient.

Although the mean power of radar transmitters is relatively small the peak power during the pulse is large, and it is the peak power which mainly determines the design of transmitting aerials. Although the result of breakdown may not be so serious in the case of a pulse transmission as for a similar c.w. power, insulation must be adequate to prevent breakdown, otherwise a considerable drop in radiated power will take place and the interference level of the

transmission will be increased owing to radiated 'brushing' noise. For the early long-wave forms of radar, peak powers of a few hundred kilowatts had to be handled by the aerials, but later the peak power was increased to the order of a megawatt, and the engineering problems of designing aerials for powers of this order are considerable.

Transmitting aerials for C.H. radar(2) were required to provide a broad horizontal polar diagram in a 'forward' direction falling off in directions at right angles and having as little as possible radiation in a 'backward' direction. It was found that best results were obtained with horizontal polarization (mainly because of the shape of aircraft) so that horizontal aerials were required. The desired
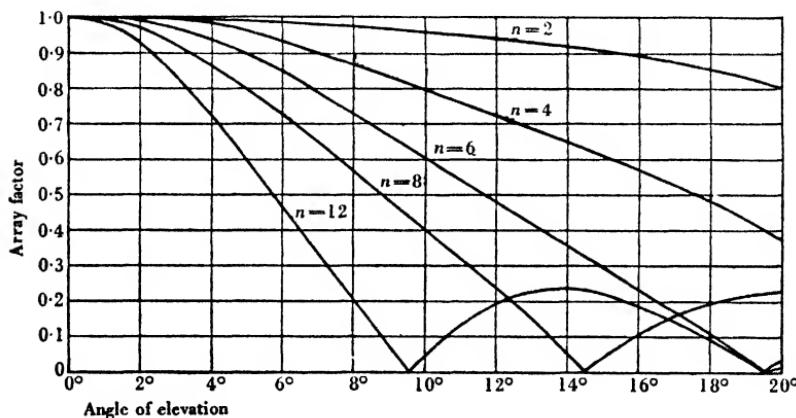
Fig. 36. Stack factors for arrays having $\frac{1}{2}\lambda$ spacing.

form of horizontal polar diagram is provided by a half-wave dipole and reflector. Several sets of these may be built into a vertical stack to provide more gain. There is a limit set to the number both by the need to illuminate most of the first 20° of elevation and by the difficulty of matching arrays with too many elements effectively in parallel.

Array factors for stacks of 2, 4, 6, 8 and 12 elements are shown in fig. 36. These have been calculated on the assumption that there are equal currents in all the elements. This will not be accurately true, particularly as regards the end-elements, but the difference in array factors produced in arrays having an even number of elements is not great. It will be seen that in order to cover

adequately 20° of elevation not more than twelve elements should be used. In addition to the effect of the array factor on the vertical polar diagram we must take into account the effect of the earth. As we have seen in § 2·6 this will produce zeroes which must be filled if we are to have complete coverage. Attempts to fill these gaps by feeding a fraction of the power into an array at different height in phase quadrature with the main array were unsuccessful. It is necessary to switch *all* the power into a subsidiary array which normally has not more than four elements so as to give effective high-angle coverage. The vertical polar diagrams obtained by use of a combination of such arrays is shown in fig. 37.
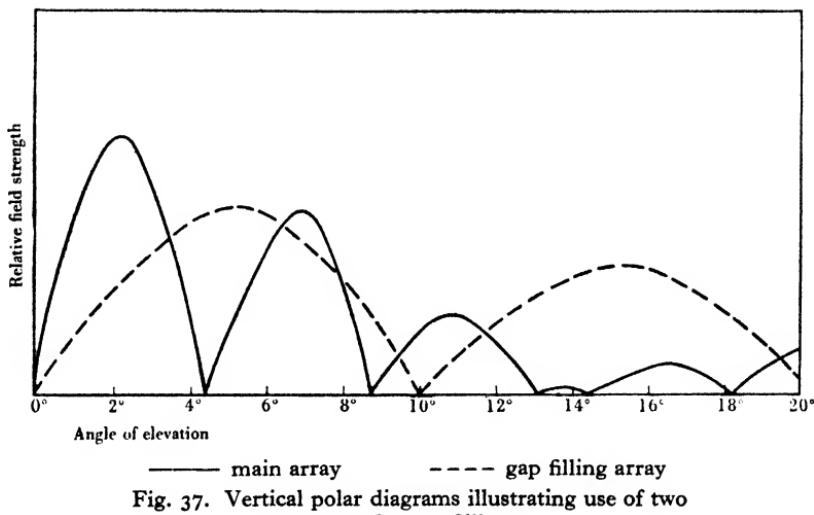
Fig. 37. Vertical polar diagrams illustrating use of two
arrays for gap filling.

Since full-wave dipoles give too narrow a polar diagram they could not be used. However, if a stack of half-wave dipoles with an even number of elements is used, each fed at one end by connecting alternately to each of the wires of an open transmission line as shown in fig. 38, a fairly well-balanced system may be made. It must be noted, however, that due to mutual interactions the dipoles do not each carry the same current, and because of this and the configuration of the connexions some unbalance must take place on the interelement feeders resulting in the emission of vertically polarized radiation. For a wave-length of 10 m. λ/d for 200 lb./mile copper wire is about 3000, so that an end-impedance of the order of

6500 ohms (twice that for a full-wave dipole) would be expected. A six-element stack would therefore have an input impedance of the order of 1100 ohms, and should be easily matched to a 600-ohm transmission line. These values were quite well borne out in practice. Since arrays had to be slung in space centre feeding was out of the question, as this can only be achieved by feeding all the elements in parallel. This would result in an input impedance so low that it could not be conveniently matched to an open-wire transmission line. We shall later see how centre feeding may be carried out by splitting the array into two or more parts. This is, however, only possible when an array is mounted on the tower itself.

The arrays used for C.H. radar(2) were backed by stacks of parasitic reflectors. From the point of view of suppression of backward radiation, however, these arrays at first were not at all successful. A considerable amount of work went into the investigation of this, as the backward radiation was exceedingly troublesome since it caused reflexions from landward objects which obscured wanted reflexions from aircraft.

6·1·1. *Effect of steel towers*

One cause of unwanted backward radiation was the excitation of the steel elements of the supporting tower. There was, of course, considerable coupling between these and the elements of the array, and it was found that quite large currents were induced in some members whose length made them nearly resonant. This had the effect of seriously distorting the polar diagram and in particular producing spurious backward radiation. An example of the polar pattern produced by a six-element stack slung between the cantilevers of a steel tower (fig. 38) is shown in fig. 39 along with the expected polar diagram. In addition, a considerable amount of vertically polarized radiation, up to 30% in field strength, was observed.

6·1·2. *Effect of interaction with supporting wires*

The second difficulty which arises with any slung array is the mutual interaction between the parasitic elements and supporting wires and tower structure. Unless the supporting wires can be broken up into short lengths (say less than $\frac{1}{10}\lambda$) by insulators this interaction affects the properties of the parasitic elements as

Main array slung between cantilevers

Inter-tower main array

Gap filler

Gap filler
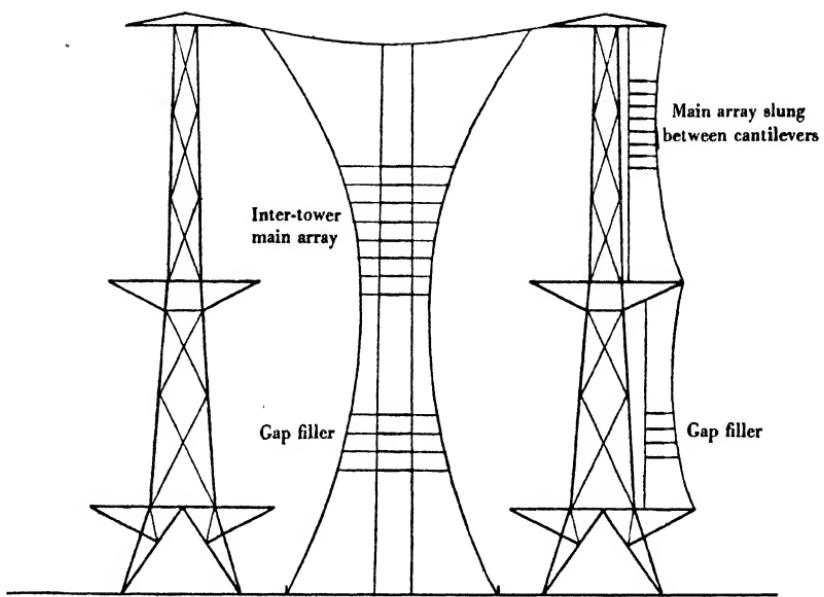
Fig. 38. Stacked arrays slung between cantilevers and between towers.

——————— Measured polar diagram

– – – – – Expected polar diagram
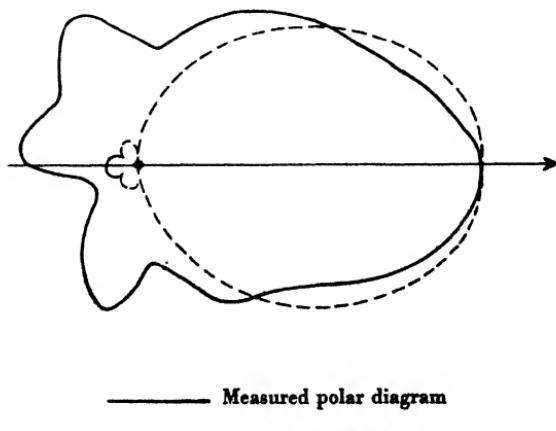
Fig. 39. Effect of steel tower on horizontal polar diagram.

reflectors. Although an optimum length for the elements for a spot frequency can be found by measuring the backward radiation for various lengths, it was found to be unsatisfactory to scale the aerial for use on neighbouring frequencies, mainly, it is thought, because of interaction with the tower structure.

6·1·3. *Inter-tower arrays*

The effect of the steel towers led to arrays being slung between pairs of towers as shown in fig. 38. To a large extent this arrangement overcomes the interaction of the array with the tower itself. The interaction with supporting wires still remains, however. The amount of vertically polarized radiation from an inter-tower array of horizontal dipoles is, however, quite small and it is possible to adjust the parasitic elements to give quite a good front-to-back ratio.

6·2. Measurements on insulators and transmission lines at high r.f. pulse powers

Before the design of a high-power transmitting aerial can be successfully undertaken, an estimate must be made of the voltages expected at various points in order to insure that adequate provision is made against breakdown or brushing.

The r.m.s. voltage V occurring between two transmission lines may be written as

$$V^2 = V_{\text{max.}}^2 \sin^2 \theta + V_{\text{min.}}^2 \cos^2 \theta, \tag{1}$$

where θ is the 'electrical distance' $2\pi x/\lambda$ corresponding to a distance x from a voltage minimum, and $V_{\text{max.}}$ and $V_{\text{min.}}$ are the maximum and minimum voltages. If a standing wave ratio S exists on the line then $V_{\text{max.}}/V_{\text{min.}} = S$. The impedance at $V_{\text{max.}}$ will be SZ_0, where Z_0 is the characteristic impedance of the line. The power P being transmitted down the line is then $V_{\text{max.}}^2/SZ_0$, or

$$V_{\text{max.}} = S^{\frac{1}{2}}(PZ_0)^{\frac{1}{2}}. \tag{2}$$

For any given load the electrical length θ between the load and voltage maximum may be calculated from standard transmission-line theory and the voltage at the load calculated. For example, if we have a load matched to a line by a stub as shown in fig. 40, on CD $S = 1$ and on AC S is determined by the impedance of the load. The electrical distance of the stub from the voltage maximum B is given by[3]

$$\cot \theta = S^{\frac{1}{2}}. \tag{3}$$

The voltage V_0 at the stub and hence on CD is then

$$V_0 = V_{\mathrm{max.}}/S^{\frac{1}{2}}. \tag{4}$$

The power input P is clearly V_0^2/Z_0 or

$$V_0 = (PZ_0)^{\frac{1}{2}}, \quad V_{\mathrm{max.}} = (PZ_0)^{\frac{1}{2}} S^{\frac{1}{2}}. \tag{5}$$

If the load is purely resistive and of value R greater than Z_0, $V_{\mathrm{max.}}$ will occur at A and the voltage at the load will be clearly $(R/Z_0)^{\frac{1}{2}} V_0$, since $S = R/Z_0$. If R is less than Z_0 the same formula applies since
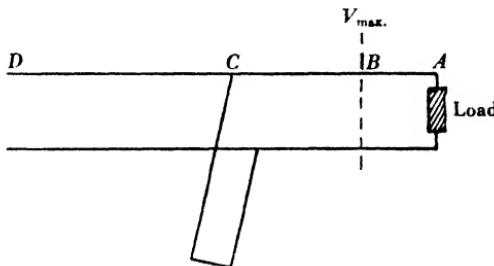
Fig. 40. Load matched to line by stub.
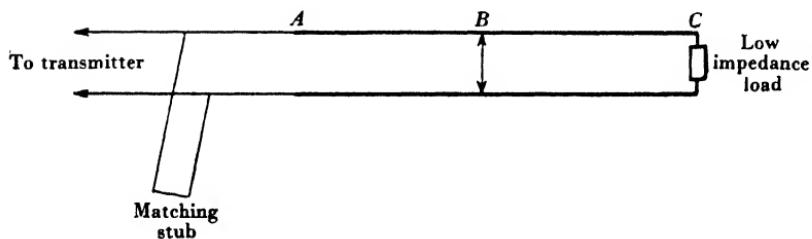
Fig. 41. Arrangement for obtaining high r.f. voltage.

S is now equal to Z_0/R. If the load has a reactive component the voltage at the load will be intermediate between $V_{\mathrm{max.}}$ and $V_{\mathrm{min.}}$ and must be determined from (1).

High voltages may be produced for test purposes by feeding a transmitter into a load consisting of a mismatched line matched to the transmitter by means of a stub. The line is terminated by a low-impedance load, the high-voltage point occurring nearly $\frac{1}{4}\lambda$ from the end. The arrangement is shown in fig. 41. The higher the standing wave ratio produced on AC the higher the voltage at B, according to equation (2). A limit is set, however, by the ability to

PLATE I
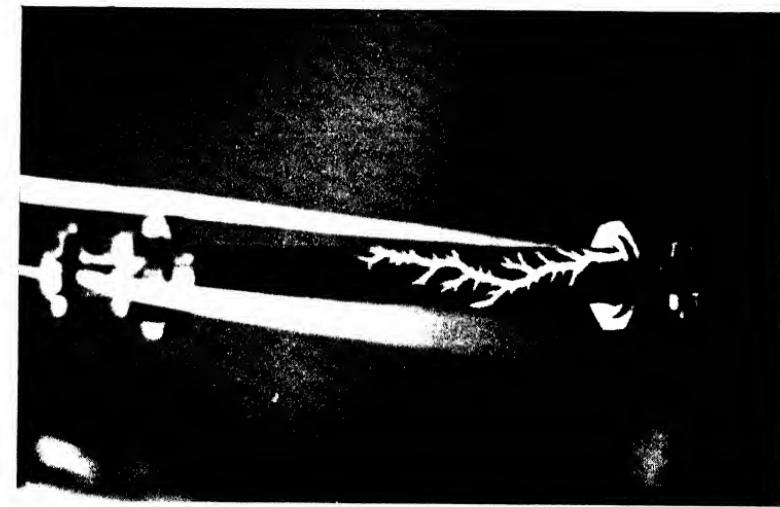
Fig. 43. Breakdown of insulator by high r.f. voltage.
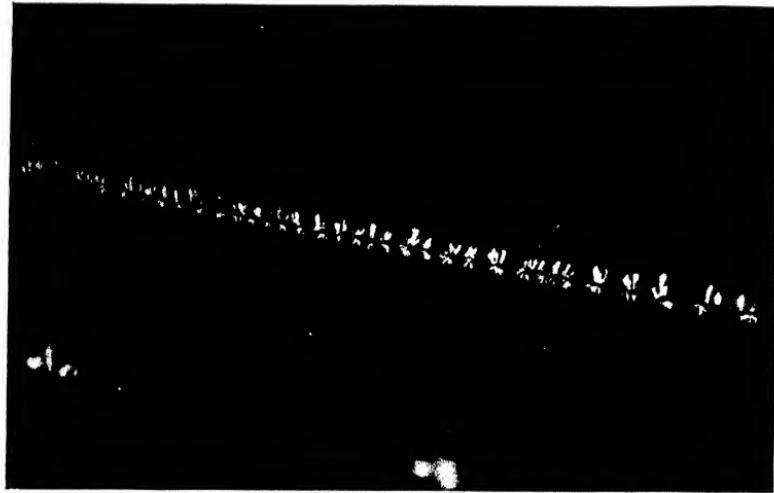
Fig. 42. Brushing from transmission lines at high r.f. voltage.

match to the transmitter by means of the stub. If a load is presented to the transmitter which is badly matched to it the power fed to the load from the transmitter will, of course, be reduced. By means of the arrangement shown in fig. 41 a standing wave ratio of 15 may be obtained on a 600-ohm line and can be matched to a transmitter. If an input pulse power of 200 kW. is available then $V_0 = 11$ kV. (r.m.s.). The maximum voltage obtainable between the lines is therefore of the order of 40 kV. (r.m.s.). The various lines under test are inserted to make up the line AC, and the power input from the transmitter is increased till sparking occurs. This normally begins as a corona discharge into the air from each wire. With the spacing between lines used in practice this takes place before breakdown between the lines. The form of the discharge will be seen from the photograph in fig. 42 (Pl. I). The maximum voltage measured between the lines which transmission lines made of wires of various gauges will support without brushing is shown in table 7.* These were obtained with wires spaced 9 in. apart, but it is found that the breakdown voltage is independent of spacing down to a few inches. These values are for pulse transmissions. The breakdown voltage, however, appears to be largely independent of pulse length for pulses varying between 5 and 50 μsec.

Table 7. *R.m.s. voltage at which corona begins on parallel copper wires spaced 9 in. apart*

Maximum r.m.s. line voltage (kV.)	Type of wire (lb./mile copper wire)
28	200
32	300
38	400
> 42	600

In order to test the breakdown properties of insulators these are placed across the lines at the high-voltage point B and the power input from the transmitter increased till breakdown takes place. This generally starts with brushing from the metal parts or 'eyes' of the insulator especially where they were in contact with the lines. With further increase in power input flash-over frequently takes place along the body of the insulator. An example of this is shown in fig. 43 (Pl. I).

* Unpublished measurements by the author and W. T. Blackband.

For slung arrays it is customary to use straining insulators consisting of rods of ceramic with 'eyes' at the end through which a loop of wire is passed. In order to test these the feeder lines are passed through the eyes and are bound firmly to the insulator. Sparking takes place at the contact between wire and insulator when the voltage between the lines is greater than about 10 kV. r.m.s. In some cases the 'eyes' are silvered in an apparent attempt to improve this contact and avoid sparking. It is found, however, that

(a) 9 in. spacing or straining insulator

(b) Same as (a) with copper corona cap

(c) 12 in. spacing or straining insulator with 3 in corona ring
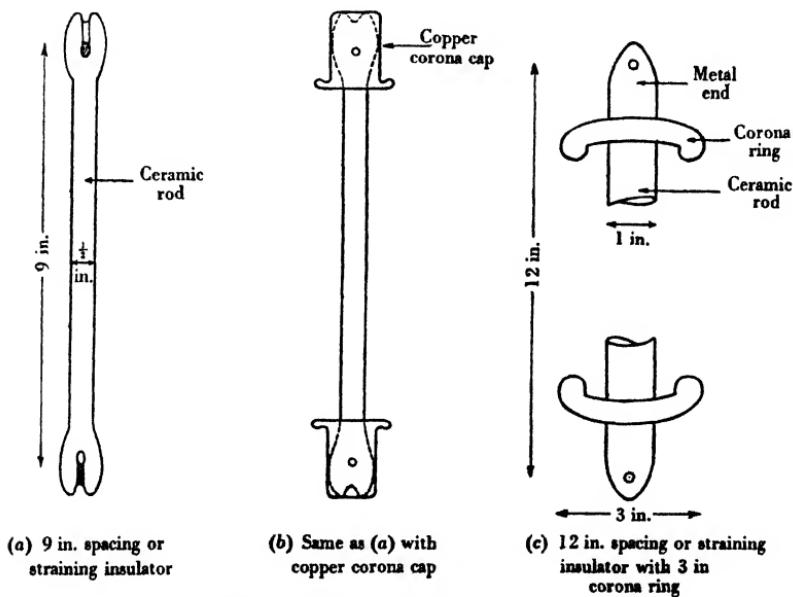
Fig. 44. Insulators and corona rings.

the silvered insulators are frequently worse from this point of view, as sparking takes place from the edge of the silvering at an even lower voltage, particularly if it has become loosened from the body of the insulator. It is quite clear, therefore, that if this type of insulator is to be used for powers of the order of 500 kW. and more some form of corona ring must be used. Examples of two types of corona ring are shown in fig. 44 as well as the bare insulator. These insulators are used for spacing transmission lines as well as straining aerials. Some figures for breakdown voltage of these insulators are given in table 8.

Table 8. *Breakdown voltage of insulators*

Type of insulator	Breakdown voltage (r.m.s.)
Fig. 44 (a)	Less than 10 kV.
Fig. 44 (b)	38 kV.
Fig. 44 (c)	Greater than 42 kV.

The addition of a corona ring is found greatly to increase the voltage at which the insulators show signs of sparking or breakdown. The corona rings, however, increase the capacity of the insulators. When placed across a transmission line these act as lumped impedances. When used to support parasitic reflectors such insulators have a marked effect on the optimum length (see Chapter 5). Some examples of the capacities of insulators are given in table 9. These capacities may be measured by means of an impedance bridge. The bridge is balanced when connected to an open line. The insulator is then placed across the line at a high impedance point, and the equivalent length of line found which just balances out the capacity of the insulator. Knowing the characteristic impedance of the line the capacity of the insulator is easily calculated.

Table 9. *Capacities of various insulators*

	Capacity (pF.)
Fig. 44 (a)	0·20
Fig. 44 (b)	0·53
Fig. 44 (c)	1·14

6·3. Effect of insulator capacity on aerial design

When the impedance of the insulators shunted across a line is appreciable compared with the characteristic impedance they have a marked effect on the phase velocity of propagation down the line. In particular, if lengths of feeder have to be accurately adjusted so as to maintain the relative phase between elements of an aerial array this has to be taken into account. For example, in stacked arrays it is frequently desired to have a length of line equivalent to half a wave-length between elements, as in the aerials shown in fig. 38. When the capacity of the spacing insulators is inappreciable the length of line required is exactly half a wave-length. Some doubt

has been cast on this point from time to time, but experiments[*] have shown that between 25 and 50 Mcyc./sec. the velocity of propagation along open-wire transmission lines, even when weathered, is 0.997 ± 0.003 time the velocity of light in free space. The capacity of the insulators, however, has the effect of reducing the length below $\frac{1}{2}\lambda$.

Consider, for example, a half-wave line with spacing insulators at its points of quadri-section as shown in fig. 45. This may be regarded as being made up of four symmetrical structures as shown in fig. 46, leaving capacities $\frac{1}{2}C$ over at each end. If several such half-wave lengths are placed end-to-end the odd $\frac{1}{2}C$ will only be left over at the ends of the whole line thus formed and will result in a shunt impedance corresponding to a capacity C.
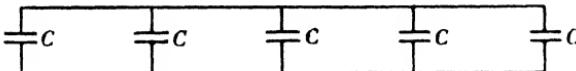
Fig. 45. Half-wave line with spacing insulators.
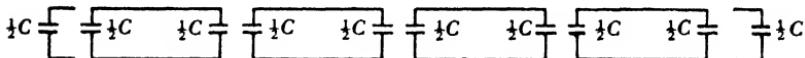
Fig. 46. Break up into symmetrical structures.

It is well known that any symmetrical network may be replaced by an equivalent length of uniform line. Let Z_0 be the characteristic impedance of the original transmission line and 4θ the total 'electrical length' of the actual lines making up the half-wave section. Let Z_1 be the characteristic impedance of the equivalent line and $4\theta_1$ its electrical length. We wish $4\theta_1$ to be equal to π so that the length plus insulators is equivalent to a half-wave line of characteristic impedance Z_1. Let α be the equivalent length of open line (Z_0) corresponding to half each insulator, i.e.

$$Z_0 \omega C = 2 \tan \alpha. \tag{6}$$

Consider a single element of fig. 46 having a value of θ_1 equal to $45°$. By a well-known result of transmission-line theory [3]

$$Z_1^2 = Z_{sc} Z_{oc}, \tag{7}$$

$$\tan^2 \theta_1 = -Z_{sc}/Z_{oc}, \tag{8}$$

* Unpublished measurements by R. J. Cary.

where Z_{oc} and Z_{sc} are the impedances seen at one end of the element when the other is open- or short-circuited. An elementary calculation gives

$$\frac{1}{Z_{oc}} = j\left[\frac{1}{Z_0 \cot\alpha} + \frac{1}{Z_0 \cot(\alpha+\theta)}\right], \qquad (9)$$

$$\frac{1}{Z_{sc}} = j\left[\frac{1}{Z_0 \cot\alpha} + \frac{1}{Z_0 \tan\theta}\right]. \qquad (10)$$

Hence $\qquad \tan^2\theta_1 = [\tan\alpha + \tan(\theta+\alpha)]/(\cot\theta - \tan\alpha), \qquad (11)$

$$Z_0^2/Z_1^2 = [\tan\alpha + \tan(\theta+\alpha)][\cot\theta - \tan\alpha]. \qquad (12)$$

Putting $\theta_1 = 45°$ we have therefore

$$\cot\theta + \tan\theta(2\tan^2\alpha - 1) - 4\tan\alpha = 0. \qquad (13)$$

As an example consider insulators across a 600-ohm line, each having $C = 2 \cdot 5$ pF. For a frequency of 53 Mcyc./sec. $\tan\alpha = 0 \cdot 25$, i.e. the insulators are each equivalent to a shunt reactance $-j \cdot 2000$ ohms. The equation for θ then becomes

$$8\cot^2\theta - 8\cot\theta - 7 = 0,$$

and the appropriate solution is $\cot\theta = 1 \cdot 56$ or $\theta = 32 \cdot 6°$.

The total length of line in the half-wave section is $0 \cdot 36\lambda$. The reduction is therefore considerable. The characteristic impedance of the equivalent line may be readily calculated from (12) and comes out as 460 ohms. With such a line it will be physically impossible to space the dipoles $\frac{1}{2}\lambda$ apart in space unless an interdipole line equivalent to a full wave is used. This would have actual length $0 \cdot 72\lambda$, and it would be difficult to take up the slack. The use of a physical spacing of $0 \cdot 36\lambda$ somewhat reduces the gain of the array and affects the array factor, but is the simplest procedure from a mechanical point of view. However, this illustrates the design difficulties introduced by the high capacity of corona rings.

The effect of the insulators on the impedance of the elements has already been discussed in Chapter 5. Parasitic reflector elements are much the most sensitive, and care must be taken to ensure that the presence of the insulators is taken into account in determining their lengths.

6·4. Design of high-power transmitting aerials

The experimental procedure which is preliminary to final design of high-power transmitting aerials is roughly as follows. The wire

required for the elements and feeders is determined by high-voltage tests. Suitable insulators are then found and subjected to high-voltage tests and their capacities determined. Experiments are then conducted to determine the optimum length of the elements. From data on the reduction of lengths of reflectors by insulators of known capacity and the curves given in Chapter 4 a first approximation to the lengths of the driven elements and parasites is made. As we have seen, a two-tier array behaves very like one having a greater number of elements as regards the action of reflectors. An experimental two-tier array is then erected and the optimum value for the length of the reflectors is found. The driven elements are then adjusted so as to be approximately resistive, though this need not be done accurately as the array will be finally matched. After the capacities of the insulators have been measured the interdipole feeders may be determined by calculation, but are usually checked by means of an impedance bridge. Finally, a prototype array is constructed and tested, the reflectors being finally adjusted for maximum reduction of backward radiation. This final adjustment is usually found to be small, but is necessary to take into account interaction with supporting wires, etc. Fortunately it is found, as we have seen in Chapter 5, that insulators behave like a fixed length of wire when attached to an aerial. A design for a fixed frequency may therefore be adapted for use at neighbouring frequencies. When the electrical design has been completed the mechanical stressing and supporting of the array must be carried out.

6·5. Switching of high-power aerials

It is frequently necessary to be able to switch the power from one aerial into another, for example, to fill gaps in the vertical coverage. Such switching has to be done by remote control, the contacts being actuated by means of a relay. To achieve the desired operation a switch must present either zero or infinite impedance when placed across a transmission line. Practical switches, however, present an open-circuit capacity and closed-circuit inductance which are by no means negligible. If, however, the switch is attached to a length of transmission line either the open or closed circuit condition may be readily compensated by a small change in the length of line. Suppose the switch is required to cause a short circuit when closed. Then if it is attached to a length of line somewhat less than $\frac{1}{2}\lambda$ long

this condition may be readily obtained. However, when the switch is opened, its residual capacity, and the reduction which has been made in line length, will result in an impedance other than the desired open circuit at the remote end of the line. Since it is frequently required that the line should be equivalent to an open circuit it is of interest to examine the conditions required for this purpose. Clearly if the switch is to produce *both* open- and closed-circuit conditions at the end of a line it must be equivalent to a length of line having characteristic impedance Z_0 equal to that of the actual line used. Let Z_{oc} and Z_{sc} be the open- and short-circuit impedances of the switch, then we must have (as in equation (7))

$$Z_0^2 = Z_{sc}/Z_{oc}. \tag{14}$$

This condition will not normally hold, since the closed-circuit impedance is in general too small. The switch may, however, be loaded with coils to achieve the required product of open- and short-circuit impedances (14). Let Z'_{sc} and Z'_{oc} be the values for the loaded switch. Then the switch is equivalent to a piece of line of characteristic impedance Z_0 of length l given by (cf. (8))

$$\tan^2\left(\frac{2\pi l}{\lambda}\right) = -\frac{Z'_{sc}}{Z'_{oc}}. \tag{15}$$

If the switch is now attached to a line of length $(\frac{1}{2}\lambda - l)$ it will produce an open circuit when open and a closed circuit when closed. If a line of length $(\frac{1}{4}\lambda - l)$ is used the reverse will be the case.

An arrangement for switching power between two aerials is shown in fig. 47. In the normal condition switches S_1 and S_2 are unenergized and are open. S_2 produces a closed circuit at A, and since $AB = \frac{1}{4}\lambda$ this produces an open circuit at B so that no power flows into the branch AB and no mismatch is produced at B. The switch S_2 produces an open circuit at C and has no effect on transmission down the branch BC or on the impedance at B. On closing both switches the conditions are reversed and all the power flows into the branch BA.

Since, when a switch is open, the whole line voltage appears across it, care has to be taken with the insulation of switches designed for high-power operation, otherwise sparking will take place from the contacts. The voltage appearing at the switch may, however, be reduced considerably by using half-wave lines for both switches

and making these of two quarter-wave sections of different characteristic impedance. From a switching point of view this behaves like. an ordinary half-wave line, but the voltage at the switch on open circuit is reduced in the ratio Z_1/Z_2, where Z_1 is the characteristic impedance of the quarter-wave section (line + switch) attached to the switch and Z_2 that of the quarter-wave section (exactly $\frac{1}{4}\lambda$) attached to the feeders.
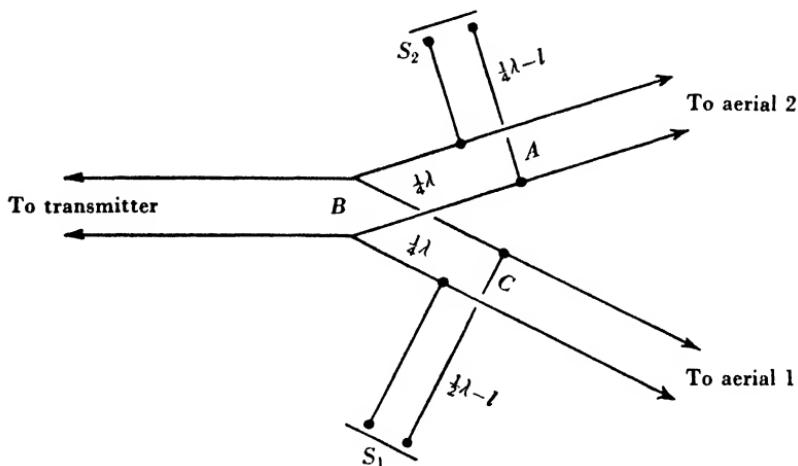
Fig. 47. Arrangement for power switching using lines and switches.

6·5·1. *Quarter-wave and half-wave stubs*

It is sometimes necessary for supporting or straining purposes to attach a stub to the feeders of such a length that it provides a perfect open circuit. A closed $\frac{1}{4}\lambda$ stub or open $\frac{1}{2}\lambda$ stub is normally used for this function. In general, the closed stub is to be preferred as its remote end is more readily supported and provides a d.c. earth for the aerial system—a useful safety measure. The physical length of a $\frac{1}{4}\lambda$ stub is not exactly half a wave-length, as it is usually closed by simply shorting with a wire of the same diameter as the lines themselves. This wire has inductance. It is found that by making the length $\frac{1}{4}\lambda - 8$ cm. a very good open circuit is obtained with standard 600-ohm line (200 lb./mile copper wire spaced 9 in.) over a band of frequencies extending from 20 Mcyc./sec. to about 60 Mcyc./sec. A closed stub of length $\frac{1}{2}\lambda - 8$ cm. gives a very good short circuit.

6·6. Switching of reflectors

It is sometimes desirable to reverse the 'forward' and 'backward' directions of an aerial. The simplest way to achieve this is undoubtedly to have two exciter curtains, one on each side of an aperiodic reflector curtain, and to switch the power between the arrays. If parasitic reflectors are used the presence of the inoperative exciter curtain tends to upset their action. However, when the space available for mounting the arrays is limited, for example, in the case of aerials for mobile use, it is extremely desirable to use only two stacks, one acting as exciter and the other as reflector and to reverse
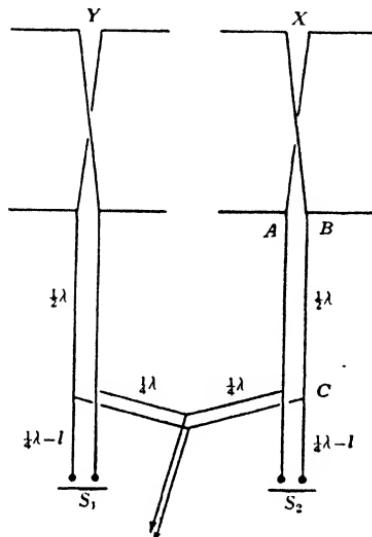
Fig. 48. Dual use of array as exciter or reflector.

their role. This may be neatly achieved as follows, and works very well for stacks having only a few elements. The arrays marked X and Y in fig. 48 are to be alternately exciters and reflectors. They are separated in the figure for convenience. The switches S_1 and S_2 are line switches capable of producing either a perfect closed circuit or perfect open circuit. If S_2 is open and S_1 closed all the power will flow to the array Y. If the distance CB is exactly $\frac{1}{2}\lambda$ a closed circuit will be presented across AB, and if the elements are of the right length the array X will act as a reflector. On closing S_2 and opening S_1 the conditions are reversed, X becoming the exciter and Y the reflector.

6·7. Aperiodic reflector curtains

Aperiodic reflector curtains as described in Chapter 5 are not normally used for wave-lengths longer than about 7 m. The number of wires required to give adequate front-to-back ratio makes them prohibitively heavy for slinging in space. For wave-lengths less than about 7 m. it is possible to erect aperiodic reflector curtains of 'chicken netting' on the face of a steel tower, and these form not only efficient reflectors but screen the aerials from the tower itself so that energy is not lost in exciting the tower structure; also the aerials gave the expected horizontal polar diagram.

6·8. Supported aerials

The use of slung arrays, as we have seen, introduces considerable difficulties in design due to the necessity of having insulators at high-voltage points. If, however, arrays are supported on a rigid structure the dipoles may be fixed near their centre at points of relatively low voltage leaving the high-voltage ends free. This not only eases the insulation problem but removes solid dielectric from the high-voltage points and so decreases the inevitable loss incurred in the dielectric under such circumstances. Undoubtedly, if it is at all possible, high-power aerials should be mounted on a wooden tower. If only a steel tower is available they should be screened from the tower by wire mesh. These measures are, unfortunately, impracticable for wave-lengths greater than about 7 m. when slung arrays must be used. The design of really high-power aerial arrays is so much simpler when these are supported that we must pause to consider the main features of such arrays.

Half-wave dipoles are supported as shown in fig. 49 by 'stand-off' insulators fixed to a wooden support. They may be either broken at the centre and fed there as shown in fig. 49 or may be continuous and fed at the end, no insulator being used at the point of attachment of the feeder. As we have seen, however, end feeding of half-wave dipoles in an array leads to unbalance in the feeders and centre feeding has many advantages. For example, the points of attachment to the dipoles are low-voltage points, and it can be arranged that the high-voltage points midway between occur where no insulators are present. Moreover, if desired, copper rods may be used for interdipole feeders so as to remove some of the troubles due to sparking.

For slung arrays it is almost essential for reasons of mechanical convenience to feed all the dipoles in parallel as shown in fig. 38, and so one is constrained to use end feeding, as centre feeding would result in too low an impedance to match to a practical line. For supported arrays, however, end feeding is no longer necessary, and the array may be split in two and fed as illustrated in fig. 50 which shows how a six-element array may be centre fed without introducing a standing wave ratio greater than 3 to 1 on a 600-ohm feeder. The impedance at *A* and *B* will be of the order of 25 ohms. The feeders between *A* and *B* may be made of copper tubing having a characteristic impedance of, say, 300 ohms. The impedance at *C* when the transformed impedances at *A* and *B* are in parallel will be
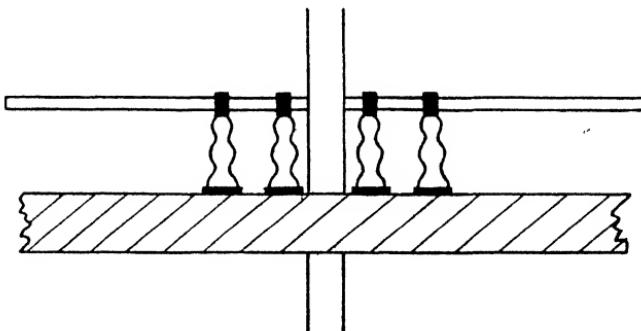
Fig. 49. Method of supporting half-wave dipole.

1800 ohms, and this may be matched readily to a 600-ohm line as it produces a standing wave ratio of only 3 to 1. Clearly considerably more than six elements could be treated in this way. In fact twelve elements would introduce only a ratio of about 6 to 1, and if desired the array could be further divided into sections and the ratio kept quite low.

6·9. Aerials for vertical polarization

For early warning radar, as we have seen, horizontal polarization was always used. For radar navigational aids such as Gee(4) and most communication purposes, vertical polarization is used. The main reason is that an 'all-round' polar diagram in the horizontal plane is more readily obtainable, and this is important from the point of view of reception in aircraft. Since some backward radiation is no longer a menace, but indeed a necessity, the acute problem of efficient reflectors no longer arises. Aerials therefore

consist usually of single vertical stacks of up to eight half-wave dipoles supported on the front of a steel or wooden tower or steel guyed mast. In order to increase the band-width properties of the aerials branch feeding is frequently adopted as shown in fig. 51. This has the advantage that all the elements are fed in phase even when the frequency is considerably changed. Moreover, $\frac{1}{4}\lambda$ transformers may be inserted to keep the standing wave ratio low. This
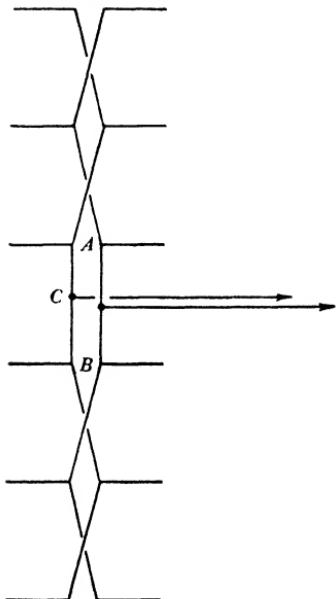
Fig. 50. Six-element centre-fed array of half-wave dipoles.
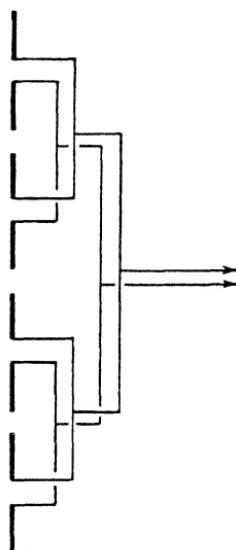
Fig. 51. Branch feeding four-element array.

arrangement, however, requires considerable space on the tower for feeders, and care must be taken in the case of a steel tower that these do not run too close to members, otherwise serious impedance discontinuities may result. When very wide-band aerials are used much more elaborate precautions must be taken. The whole subject of wide-band aerials has, however, been reserved for a separate chapter.

REFERENCES

(1) Smith, R. A. *Electronics*, Chapter VI (Radar). Pilot Press (1947).
(2) Watson-Watt, R. A. The evolution of radar. *J. Instn Elect. Engrs*, **93**, IIIA, 11 (1946).
(3) Jackson, Willis. *High-frequency Transmission Lines*. Methuen (1945).
(4) Smith, R. A. *Radio Aids to Navigation*, Chapter 6. Cambridge University Press (1947).

Chapter 7

RECEIVING AERIALS FOR WAVE-LENGTHS OF A FEW METRES

7·0. Introduction

Receiving aerials may perform two very distinct functions. They may simply be the means of picking up radio-frequency energy or they may be the vital measuring part of the equipment to which they belong, the receiver merely acting as an indicating instrument. An example of the former use is the ground-receiving aerials for the Gee locking signals[1], and, of the latter, the receiving aerials for C.H. radar[2]. For the former use, only moderate care is required to make the aerials efficient, but for the latter use, of course, the utmost care is required, otherwise the value of the whole equipment may be seriously reduced through the aerials giving erroneous information.

7·1. Receiving aerials for vertical polarization

An aerial giving wide cover in the horizontal plane is frequently required. In order to increase the gain, as much vertical stacking as practically possible is used, since the signals generally come from a very small angle of elevation. Stacks of vertical half-wave dipoles are used exactly as for transmitting (Chapter 6). The requirement of insulation against high-voltage breakdown now no longer exists. The aerials for reception and transmission only differ as regards feeding arrangements. The high voltages developed in the feeders in the case of transmission makes the use of open-wire balanced feeders desirable. For reception, however, a single concentric feeder has many advantages. The chief is that it is much less liable to pick up undesired signals such as those due to ignition interference from passing motor vehicles, etc. Open-wire lines are usually subject to some unbalance and may pick up spurious signals more readily than concentric lines. Concentric lines require many more spacing insulators and have a higher loss than open lines unless they are made of large diameter tubing when they become very expensive. If a very long feeder run is required the attenuation of the concentric line must be considered.

With stacked arrays it is rather inconvenient to use unipoles and dipoles are normally employed, a balance-to-unbalance transformation, taking place at the foot of the array. If a single unipole is sufficient, however, it is a simple matter to feed it directly from a concentric feeder as shown in fig. 52. This arrangement is widely used for communications purposes. The only difficulty is in providing a suitable earth for the outer conductor of the concentric line. Three arrangements are shown in fig. 52. In (a) a metal sheet earth is used. Theoretically this should be infinite in extent, but if greater than $\frac{1}{2}\lambda$ in radius is fairly effective. A better and more convenient arrangement is shown at (b) where use is made of the
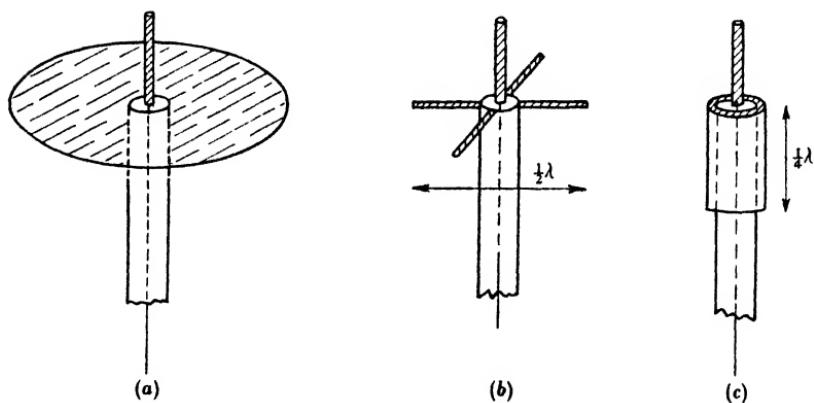
Fig. 52. Three arrangements for providing an earth
for feeding a unipole.

half-wave dipole earth. The centre point of a half-wave dipole is at a very low impedance point with respect to free space, and when connected as shown makes quite an effective earth. The arrangement is not critical so long as fairly stout rods ($\frac{1}{4}$ in. diameter or more) are used for the earthing elements. One such element is sufficient, but two are frequently used at right angles for purposes of symmetry. The arrangement shown at (c) makes use of the fact that the $\frac{1}{4}\lambda$ sheath over the outer conductor of the concentric line forms a concentric line with the outer and produces a very low impedance at the top end of the line. It should be noted that currents flow in the outer surface of the sheath, and the unipole and sheath together behave rather like a $\frac{1}{2}\lambda$ dipole. In each case, the top of the concentric line is closed with insulating material, which produces in effect

a lumped impedance across the line at the foot of the unipole. It is convenient to use line having a characteristic impedance of about 35 ohms so as to match the unipole, any residual reactance being removed by adjustment of the length of the unipole. If such feeder is not available a $\frac{1}{4}\lambda$ transformer must be used, the unipole plus termination again being adjusted so as to be purely resistive.

7·1·1. *Balance-to-unbalance transformation*

For stacked arrays dipoles are more convenient as they may be interconnected by open-wire feeders and avoid the necessity for weatherproof terminations and feeder junctions. In this case a balance-to-unbalance transformation is required if the aerial is to be fed by a single concentric line. Several line circuits exist for
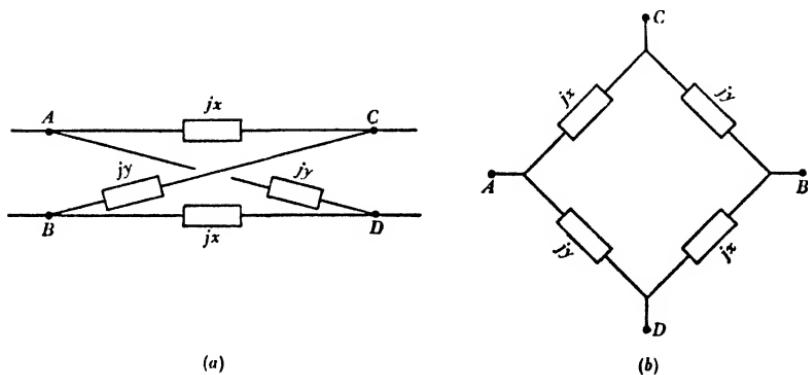
Fig. 53. Lattice network.

accomplishing this, and are sometimes used. They are, however, somewhat clumsy at wave-lengths of a few metres, and a discussion will be given later of their use with aerials for $1\frac{1}{2}$ m. The 'lattice' network provides perhaps the simplest and most compact transformation unit, and is widely employed at wave-lengths longer than 3 m. The network is shown in fig. 53 (*a*) or its equivalent form 53 (*b*). It provides an impedance transformation as well as a balance-to-unbalance transformation. Because of the importance of this network we shall give a brief description of its properties and indicate some of the precautions which must be taken in the design of such networks for wave-lengths of a few metres.

The impedances jx, jy shown in fig. 53 are considered as pure reactances either inductance or capacitive. Since the network is

symmetrical it may be regarded as equivalent to a length of transmission whose electrical length θ and characteristic impedance Z_0 are given by (cf. equations (7), (8) of Chapter 6)

$$\tan^2\theta = -Z_{sc}/Z_{oc}, \tag{1}$$

$$Z_0^2 = Z_{oc}Z_{sc}, \tag{2}$$

Z_{oc} and Z_{sc} being the open- and short-circuit impedances. Clearly we have

$$Z_{sc} = 2j\frac{xy}{x+y}, \tag{3}$$

$$Z_{oc} = \tfrac{1}{2}j(x+y), \tag{4}$$

so that $$Z_0^2 = -xy \tag{5}$$

and $$\tan^2\theta = 4xy/(x+y)^2.$$

If the network is to be equivalent to a real line then x and y must be of opposite sign. Suppose x corresponds to an inductance and y to a capacity we have, writing

$$x = \omega L, \quad y = -1/\omega C, \tag{6}$$

$$Z_0^2 = L/C, \tag{7}$$

$$\tan\theta = \frac{2\omega(LC)^{\frac{1}{2}}}{(1-\omega^2 LC)}. \tag{8}$$

We note that if $$x = -y, \quad \omega^2 = 1/LC, \tag{9}$$

then $\theta = \tfrac{1}{2}\pi$ and the network behaves like a quarter-wave transformer. When this condition holds, moreover, the network has an additional very interesting property. Consider an arrangement such as shown in fig. 54(a) or its equivalent circuit 54(b), in which a balanced load $2r$ is placed across CD, E being the electrical centre of the load. If B is earthed and an alternating voltage applied between A and earth, it is easy to show by joining E and B by a wire of zero resistance, calculating the current in it and equating this to zero, that the condition for zero voltage at E is $x = -y$ or $\omega^2 = 1/LC$, just condition (9). If this condition is in fact satisfied it is easy to see that the current distribution must be as shown in fig. 54(b). That the value of i' must be the same for branches AC, AD follows at once from the fact that the voltage drop along ACB is equal to that along ADB, this being independent of i if $x = -y$. This shows that

no current flows in EB. It follows at once that the value of i is the same in the two branches. If e is the voltage applied at A we have

$$e = jxi'. \tag{10}$$

Also

$$ri' = jyi. \tag{11}$$

Therefore

$$e = -\frac{xy}{r}i. \tag{12}$$

The input impedance Z_i is just $E/2i$, so writing $x = \omega L$, $y = -1/\omega C$, we have

$$Z_i = \frac{L}{2Cr}. \tag{13}$$
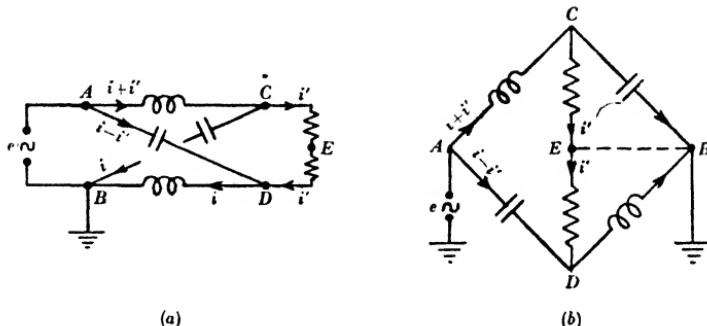
Fig. 54. Lattice network as balance-to-unbalance transformer.

If the load is a pure resistance, Z_1 will also be purely resistive. Equation (13) may have been deduced from the fact that the network behaves as a quarter-wave line with $Z_0^2 = L/C$. Any transformation, within practical limits, may be obtained by choosing L and C to satisfy (9) and (13).

In the practical design of lattice networks the required capacities and inductances are first calculated and the network is then assembled. Generally it is placed in a weatherproof box, and if this is made of metal there will be considerable capacity between the inductances and the box, and interaction between the inductances themselves, so that adjustments will require to be made. These are best made by means of a high-frequency impedance bridge, final adjustments being made to the inductances after a suitable value for the condensers has been found. The lattice network plus these stray impedances may be fairly well represented by the equivalent circuits shown in fig. 55. The capacities C' and C'' will not in general

be equal owing to the presence of the containing box which is connected to B. C'' may be removed by a compensating inductive component of the load, and C' may have to be removed by an inductance across the input or by further compensation at the load.

The band width of such a transformation unit is quite large provided the ratio $Z_i/2r$ is not large or small. The series circuits DAC, CBD have the resistance $2r$ in parallel and are highly damped if r is of the same order as ωL. When the frequency is changed so that $\omega^2 \neq 1/LC$ the perfect balance is, of course, lost so that E is no longer at earth potential. The percentage unbalance will be approximately equal to the percentage change in frequency, so that for small changes it will generally be unimportant.
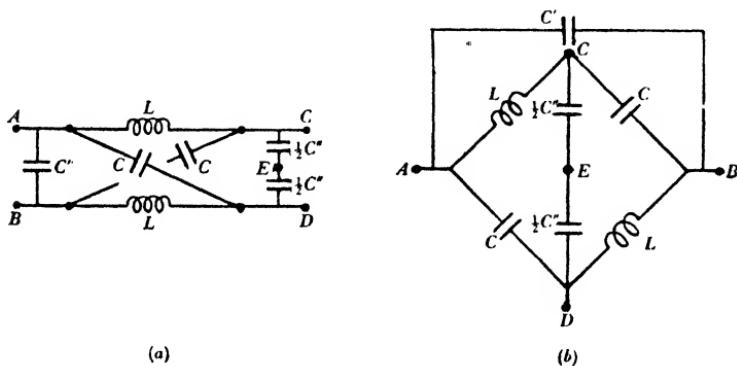
Fig. 55. Equivalent circuits of lattice network
to take account of stray capacities.

An ordinary h.f. transformer with centre-tapped secondary may, of course, be used to effect a balance-to-unbalance transformation, but the design of such transformers for frequencies over 20 Mcyc./sec. is very difficult owing to the high self-capacity of the windings. It is also necessary to introduce very good electrostatic screening between primary and secondary in order to avoid electrostatic transfer of energy. It is therefore usually preferable to use either the lattice network or some line circuit such as one of those described in Chapter 9.

7·1·2. Concentric lines for receiving aerials

The type of line to be used for an aerial system will depend on the length of feeder run and on the attenuation which can be tolerated. It is of interest to note that if the noise received in the aerial is very

much greater than that generated by the receiver, attenuation in the feeder run will not affect the signal-to-noise ratio till it reduces the aerial noise to a value comparable with receiver noise. For short runs, feeders with a solid dielectric, usually of polythene, are very convenient. They may readily be made weatherproof by means of a suitable termination. This is essential if constancy of performance is to be maintained. A great deal of effort was expended during the war in designing weatherproof terminations as the attenuation of polythene cables generally increases greatly with ingress of moisture and the phase velocity of propagation along the cable is also changed. Lead-covered cables are generally preferred for outdoor use, since braided cables become corroded, even when covered with weatherproof material, and this causes an increase in attenuation and inconstancy of performance.

For long feeder runs, air-spaced, copper concentric lines are usually preferred owing to their low attenuation. For constancy of performance these must be kept dry and airtight and are usually kept under pressure.

7·2. Receiving aerials for horizontal polarization

For horizontal polarization balanced dipoles are always used, but the same considerations regarding interference apply. Except when perfect balance is required, as in direction-finding aerials, a balance-to-unbalance transformation is used and the signals are led to the receiver by means of a single concentric line.

7·3. Receiving aerials for C.H. radar

As we have seen, when the aerials form the measuring part of an equipment, very great care is required in their design. For example, in direction-finding, the aerials largely determine the accuracy of the system. As illustrations of this type of aerial system we shall consider briefly the receiving aerial systems of C.H. radar and v.h.f. direction finding. In C.H. radar[2] the aerials were the part of the equipment used for measuring bearing and elevation, the rest of the apparatus being used as an indicator and for measurement of range. A great deal of experimental work therefore went into the design of receiving aerials. This has been described in some detail by R. A. Smith and C. Holt Smith[3], and we shall only describe briefly the essential features of the aerial system.

A great deal of work has been done on the use of goniometer technique for direction finding, and the precautions necessary to avoid errors in aerial systems are well known(4). Most of this work, however, has been concerned with vertical polarization and with highly reactive aerials. The high frequencies used for radar permit the use of tuned aerials so that exact matching to the transmission-line system is possible. This introduces new problems. The errors
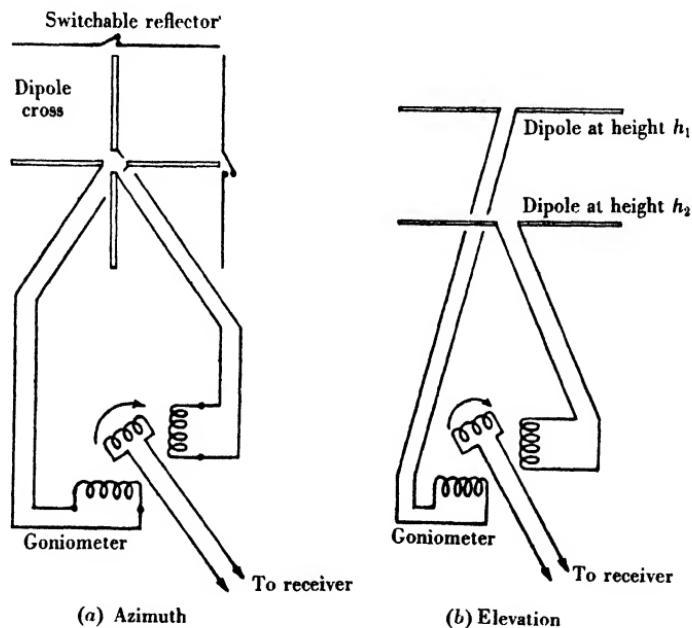
Fig. 56. Schematic arrangement of C.H. radar d.f. system.

in direction finding in the band of frequencies with which we are concerned (20–50 Mcyc./sec.) have been studied by Smith-Rose and Hopkins(5). These authors were mainly concerned with the use of vertical polarization and with relatively low aerial systems. The necessity for using high towers to increase range of detection, together with the need to conduct the received signals to a receiver room a few hundred yards from the towers, led to increased difficulties in keeping direction-finding errors small.

The essentials of the receiving system are shown in fig. 56. Signals received in two crossed dipoles are compared in a goniometer to determine the azimuth of the reradiating aircraft. Sense is deter-

mined by noting the effect on the received signal of switching a reflector aerial in and out of circuit, and elevation is measured by comparing, in the goniometer, signals from aerials at two different heights. A typical receiving array on a 240 ft. wooden tower is shown in fig. 57.
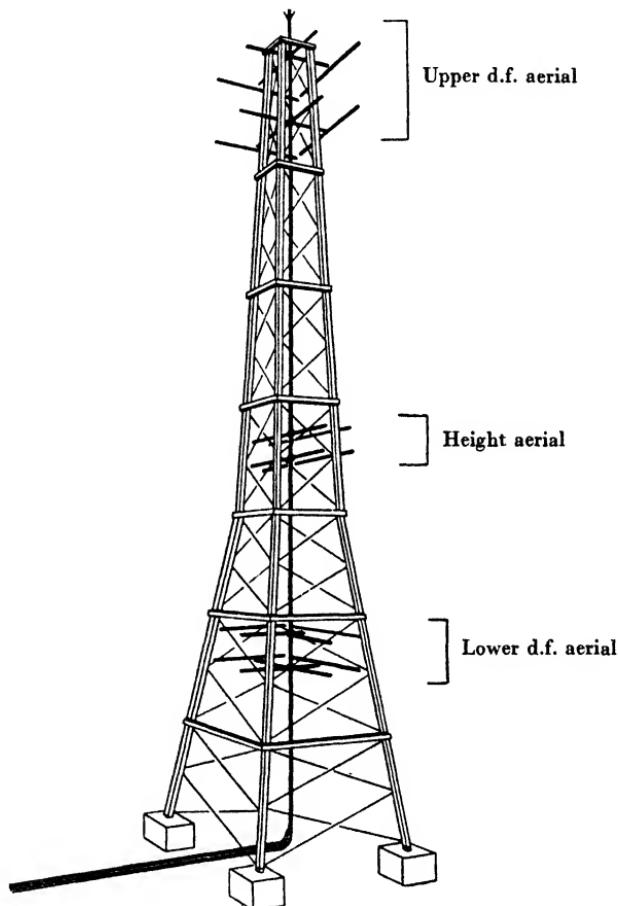
Fig. 57. Arrangement of d.f. and height-finding aerials on 240 ft. mast.

Because of the importance of perfect balance in a direction-finding system, balanced lines were used consisting of pairs of copper concentric feeders bonded together. These retain the anti-interference properties of single lines provided they are adequately bonded.

Most of the errors arising in a d.f. system are due to four causes:

(1) Variation of feeder characteristics.

(2) Interactions between dipoles and reflectors.

(3) Imperfect bonding of feeders.

(4) Sensitivity of the system to vertical polarization.

There are, in addition, certain inherent errors such as the small quadrantal error due to the departure of the polar diagram of a half-wave dipole from the $\cos\theta$ form (see Chapter 2).

Some of the errors listed above may be removed by precautions in design. Others are inherent in the system, but, if constant, may be removed by calibration. Errors due to vertical polarization cannot be removed and vary with angle of elevation. When small angles of elevation are involved and feeders, etc., are carefully balanced, the errors due to this cause can be kept small. The two most serious sources of error are feeders and reflectors. It is found that the greatest care must be taken in the design of the aerial system itself in order to eliminate these errors.

Precautions must be taken to ensure that the attenuation of corresponding pairs of feeders is small, equal and constant, and that the electrical lengths are as near as possible equivalent. When this is done, errors are mainly due to mismatches, which cause standing waves on the lines. In order to understand this, we should have to discuss in some detail the theory of the goniometer associated with long feeders. This is given fully by Smith and Holt Smith(3), who have shown that unless either the aerials are matched exactly to the feeders, or the goniometer is matched at the d.f. position, changes in electrical length of the lines will result in serious bearing errors.

If a line is matched at *either* end the phase characteristic is linear, i.e. the phase change over a length of line is proportional to the length. Moreover, the voltage produced at the end of the line depends on its length unless this condition is satisfied. Now a goniometer corresponds to a purely inductive termination when it is turned to the position of minimum signal so that one must either match carefully at the aerial end or use a resistive (and therefore lossy) matching unit at the goniometer. The latter procedure is only possible for wave-lengths greater than 10 m. where the aerial noise exceeds the receiver noise. For all wave-lengths in use for C.H. radar, however, it is essential to match the aerials to the lines if the most accurate results are required.

Vertically polarized radiation may be picked up on the feeders if they are not carefully bonded together. This was certainly a considerable source of error in early radar systems, as could be shown by removing some of the few bonds which existed, with consequent change of error. Fortunately, the towers used to support the aerials had a stout central lightning conductor, and it was possible to bond all the feeders solidly to this, approximately every quarter wavelength down the tower. The bond at the dipole positions is most important as it forms an essential part of the aerial system. It was soon realized that no repeatable impedance measurements could
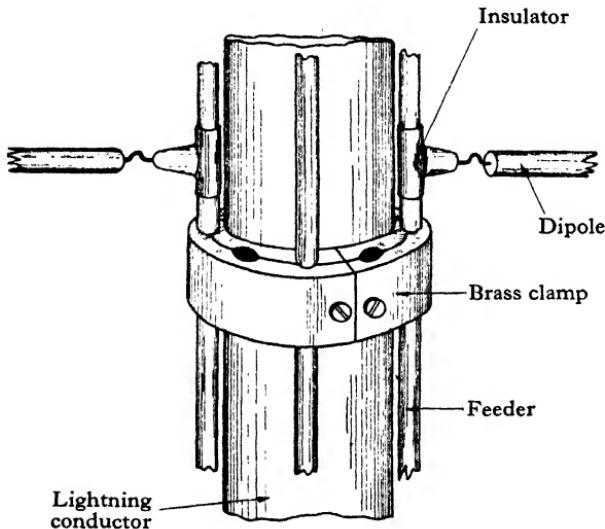
Fig. 58. Arrangement of feeders showing bonding and dipole connexions.

be made until a stout standard bond had been devised, all the feeders being clamped to a casting fitting on the lightning conductor, as shown in fig. 58.

Errors due to reflectors caused a great deal of trouble. The same considerations apply to reflector switches as for line switching of transmitting aerials (§6·5). Either a perfect open circuit or perfect closed circuit must be applied to a reflector to make it inoperative or active. This condition is by no means fulfilled by even the best switches which are available. Line switching with loaded switches is therefore introduced, but great care must be exercised in arranging the lines so as not to interact with the dipoles and reflector elements.

Again the position of the reflectors relative to the dipoles is found to be of the utmost importance. It is found necessary to increase the spacing between reflector elements and dipoles to at least 0.26λ, so that the reflectors lie well beyond the ends of the dipoles at right angles. Only by this arrangement may substantial freedom from errors be achieved.

The arrangement for switching arrays of reflectors by means of a single switch is interesting and is shown in fig. 59.

In order to obtain increased range stacked arrays are used: two-tier arrays for the 10 m. band and four-tier arrays for the 6 m. band. The principles of design of these are straightforward after the fundamental requirements have been appreciated and the fundamental measurements of impedance made, the results of which have already been discussed in Chapter 1 and are illustrated in figs. 3 and 4. The building of elements into arrays is straightforward, since good agreement is found between measured impedances and those calculated from the measured self-impedances and theoretical mutual impedances. Matching is carried out by means of stubs made of copper concentric line.
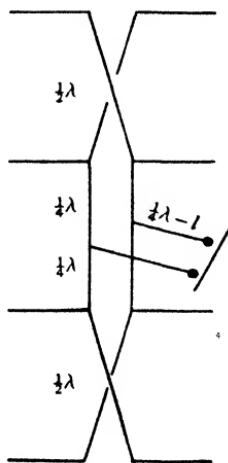
Fig. 59. Reflector switching arrangement.

A typical calibration curve of the aerials designed as described is shown at fig. 60. It will be seen that some errors still remain. These are due to inherent causes, such as the octantal error of the dipoles, combined with a residual error due to interaction between dipoles and reflector lines. The errors are, however, fairly small, so that they may be removed by calibration. If it were possible to dispense with the switched reflectors an aerial system could be made having errors very little greater than those due to inherent causes.

7·4. Aerials for v.h.f./d.f.

Direction finding on v.h.f. communications from aircraft played a large part during the war (6). This took place in the frequency band 100–150 Mcyc./sec. The fundamental problems associated with Adcock direction finders in this band had already been extensively

studied by Smith-Rose and Hopkins(5). Most of the problems arising with the aerials for v.h.f./d.f. were, however, associated with obtaining 'sense' rapidly and in a way which would not upset the accuracy of the system.

The v.h.f./d.f. system as applied to British v.h.f./r.t. was developed at the Royal Aircraft Establishment, Farnborough, as part of the v.h.f. programme. One type of aerial system which gives satisfactory results consists of an elevated H type aerial, the dipole lengths
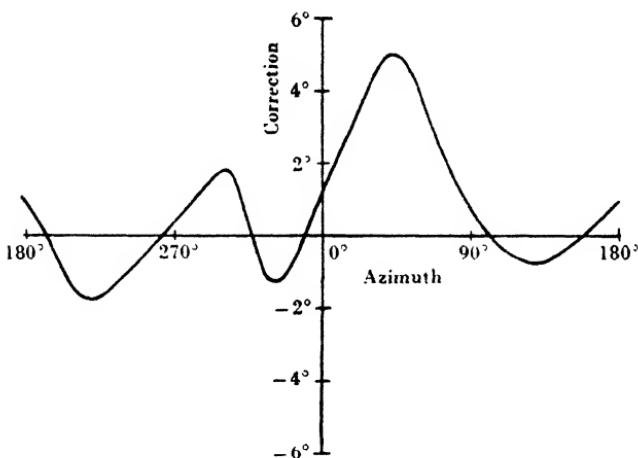
Fig. 60. Calibration curve of d.f. aerials.

and spacing being so chosen as to give the system the maximum possible sensitivity. Provision is made for rotating the aerial head by means of a metal shaft which also contains the twin feeders from the aerial system to the receiver, situated at ground-level some 16 ft. below. A balance-to-unbalance transformer is connected to the feeder line at the base of the shaft, the unbalanced side being connected to the receiver input. An aural type of indication is used. This aerial produces a 'figure of eight' pattern, so a suitable method of sense determination is still required. This is obtained by distorting the polar diagram by fixing a parasitic element in proximity to each aerial dipole. If, say, the minima of the figure of eight occur at 0° and 180°, by suitable adjustment of the length of the parasitic elements and their spacing the angular positions of the maxima may be displaced from their normal position at 90° and 270° to, say, 45° and 315°. Hence the correct bearing may be determined by

rotating the aerial head and selecting the minimum lying between adjacent maxima.

An alternative method of determining sense is as follows. Parasitic elements are again mounted in proximity to each aerial dipole, but, in this instance, they are mounted on a line at right angles to the line joining the aerials. Unlike the previous system, however, the parasitic elements are broken at the centre, the halves being connected to a wafer-type switch, which is built into the dipole insulator and which may be remotely operated. With the switches closed and the spacing and dipole lengths suitably adjusted, the polar diagram is similar to that obtained with the former aerial system. On opening the switches, however, the halves of each dipole are connected only by the capacity across the switches, and by suitably adjusting this each parasitic element acts as a director, thus reversing the polar diagram. The method of determining sense is, therefore, to swing the aerial system 10–20° off the minimum position and then operate the sense switch. The output of the receiver will then either increase or decrease, depending on the relative position of the parasitic element with respect to the aerial and the signal source. The change in audio output is of the order of 5 db. at the ends of the band rising to approximately 15 db. at the midpoint. The later type of aerial is shown in fig. 61 (Pl. II).

REFERENCES

(1) Smith, R. A. *Radio Aids to Navigation*, Chapter 6. Cambridge University Press (1947).

(2) Watson-Watt, R. A. The evolution of radar. *J. Instn Elect. Engrs*, **93**, IIIA, 11 (1946).

(3) Smith, R. A. and Holt Smith, C. Elimination of errors from crossed-dipole direction-finding systems. *J. Instn Elect. Engrs*, **93**, IIIA, 575 (1946).

(4) Keen, R. *Wireless Direction Finding*. Iliffe and Sons (1938).

(5) Smith-Rose, R. L. and Hopkins, H. G. Radio direction finding on wave-lengths between 6 and 20 metres. *J. Instn Elect. Engrs*, **83**, 87 (1938).

(6) Smith, R. A. *Radio Aids to Navigation*, Chapter 4. Cambridge University Press (1947).

PLATE II
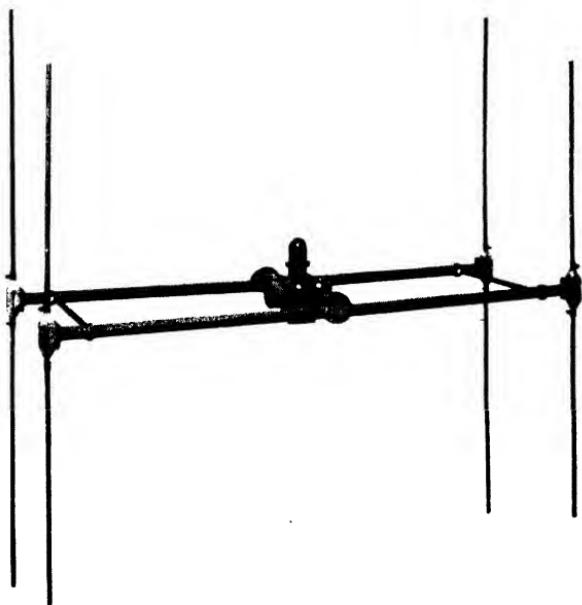
Fig. 61. V.h.f./d.f. aerials.
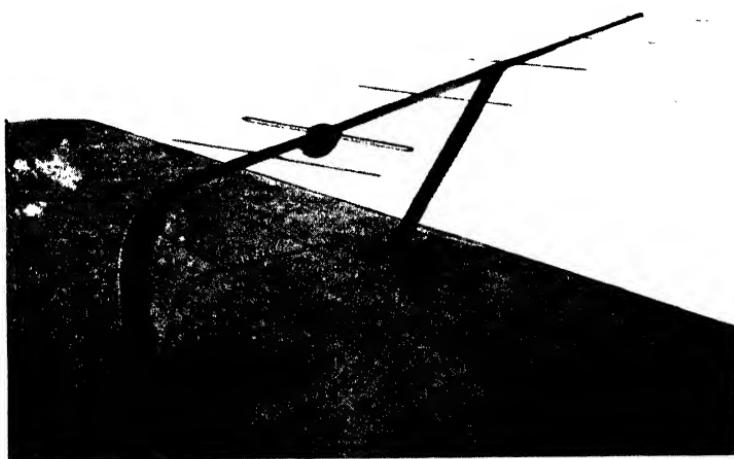
Fig. 89. Yagi homing aerial on wing of Liberator aircraft.

Chapter 8

LONG-WIRE AERIALS

8·0. Introduction

The elements of which the aerial arrays, which we have considered so far, have been built have consisted of half-wave and full-wave dipoles. There are, however, many other ways of making a broadside array using long continuous wires. A great deal of ingenuity has gone into the design of directional communication aerials using continuous wires. There are two main types of long-wire aerials. In the first the wire is bent in some arrangement which enables various parts of it to radiate in phase and other parts to cancel. An example of such aerials is the Sterba array(1). One of their chief advantages is that d.c. or low-frequency a.c. may be passed through the aerial to prevent the formation of ice in severe weather conditions. This type of array was not much used for radar purposes, mainly because broadside arrays of dipoles are simpler to construct and have better frequency characteristics. A great variety of such arrays have, however, been used for communication purposes. We shall not deal with them further here as they are mostly covered by standard text-books (e.g. Ref. (1)).

The second type of long-wire aerial which has been widely used for many years consists simply of an arrangement of long straight wires having either their ends free or terminated by a resistance to remove standing waves. The latter type of aerial is known as a travelling wave aerial. The 'diamond' or 'rhombic' aerial is the best known form. Unterminated aerials are exemplified by the V aerial.

Such aerials have been used as ground receiving aerials for radar navigational aids such as Gee, and have also been widely used for point-to-point communications. The only new points of design which have arisen during the war years have resulted from the use of shorter wave-lengths. It has thus been possible to use wires many wave-lengths long without requiring a great deal of space for the aerial. We shall consider only briefly the fundamental properties of such long-wire aerials.

8·1. The long straight wire as a radiator

The problems of transmission and reception are quite complementary in virtue of the reciprocity relations (Chapter 3), so that any conclusions arrived at for the transmitting condition apply equally well for reception. So far we have only considered a linear radiator having a length $\frac{1}{2}\lambda$ and carrying a sinusoidal current having the form of a 'standing wave'. (Note that the full-wave dipole fed at the centre is not analogous as there is a reversal in phase of the current at the centre.) We must now consider radiation from a straight wire of any length carrying a sinusoidal current. We shall consider two extreme cases:

(1) when one end of the wire is free and the standing wave ratio on it is infinite;

(2) when one end is terminated so that the standing wave ratio is unity.

We have already seen that the condition (1) is only an approximation even for a wire $\frac{1}{2}\lambda$ long. The approximation, however, is a good one for calculating the polar diagram of such an aerial. We shall first of all deal with a wire extending in one direction only from the feed point. The above approximation of choosing a sinusoidal form with exact zeroes at the nodes becomes worse as the length of the radiator is increased. When the aerial is more than three or four half-wave-lengths long experiment shows that the current at the nodes farthest from the free end is by no means zero, and finally when more than ten half-waves are included the standing-wave-ratio at the fed end becomes nearly unity. The impedance looking towards the free end tends to become resistive and the aerial behaves somewhat like a lossy line, whose input impedance is independent of its length after this has increased beyond a certain value. (There are, however, significant differences; see below.)

The case of an infinite straight wire of circular cross-section (radius a) has been treated theoretically by Schelkunoff[2], who has shown that, provided a/λ is so small that $\log_e \lambda/a$ is much greater than unity, the input impedance Z_i is given by

$$Z_i = 60(\log_e \lambda/2\pi a - \gamma), \tag{1}$$

where γ is Euler's constant, $0.577\ldots$. (Schelkunoff gives twice this value, but his calculations refer to a wire infinite in both directions and broken at the feed point.)

For very long wires, or wires correctly terminated, the travelling wave form of the current given by

$$i = i_0 e^{-2\pi j x/\lambda} \tag{2}$$

is quite a good approximation, though it would probably be improved by adding an exponential decay factor to account for loss. We may note that Schelkunoff has shown that for a perfectly conducting wire the *voltage* is not attenuated.

For a short wire of length l the 'standing wave' form given by

$$i = i_0 \sin \frac{2\pi}{\lambda} (l - x) \tag{3}$$

is a good approximation, but for intermediate values there is no simple expression available. However, by considering the properties of aerials carrying currents (2) and (3) we may obtain a very good idea of the behaviour of long-wire aerials in general. We may note, however, a tendency in the existing literature to assume that the actual current in the aerial is given by either (2) or (3) and a failure to appreciate the approximate nature of the assumed currents.

8·1·1. *Polar diagram of radiation from wire with free end*

We shall first use the form (3) to calculate the polar diagram of a wire a few wave-lengths long with a free end. By proceeding as in §3·3·2 we see that the field strength E at large distance R, radiated by a wire carrying the current (3), is given by

$$E = \frac{60\pi \sin\theta i_0}{\lambda R} \int_0^l \sin \frac{2\pi}{\lambda}(l-x) \, e^{2\pi j x \cos\theta/\lambda} \, dx, \tag{4}$$

where θ is the angle made with the wire.

When the length l is an integral number n of half wave-lengths this reduces to the simple forms, omitting phase factors,

$$E = \frac{60 i_0}{R} \frac{\cos(\pi l \cos\theta/\lambda)}{\sin\theta} \quad (n \text{ odd}) \tag{5}$$

$$= \frac{60 i_0}{R} \frac{\sin(\pi l \cos\theta/\lambda)}{\sin\theta} \quad (n \text{ even}). \tag{6}$$

The maximum of the radiation pattern only occurs at $\theta = \frac{1}{2}\pi$ for a wire $\frac{1}{2}\lambda$ long, i.e. $n = 1$. For n even there is a zero at $\theta = \frac{1}{2}\pi$, as we should expect. The positions of the first (principal) maximum and first zero of the pattern are shown in fig. 62.

8·1·2. *Polar diagram of very long or terminated wire*

When we use expression (2) for the current, which is a good approximation for a very long or terminated wire, we have for the radiated field strength

$$E = \frac{60\pi \sin\theta i_0}{\lambda R} \int_0^l \exp\left[-j\frac{2\pi x}{\lambda}(1 - \cos\theta)\right] dx, \tag{7}$$

which, omitting the phase term, gives

$$|E| = \frac{60 i_0}{R} \cot\frac{\theta}{2} \sin\left[\frac{2\pi l}{\lambda}\sin^2\frac{\theta}{2}\right]. \tag{8}$$
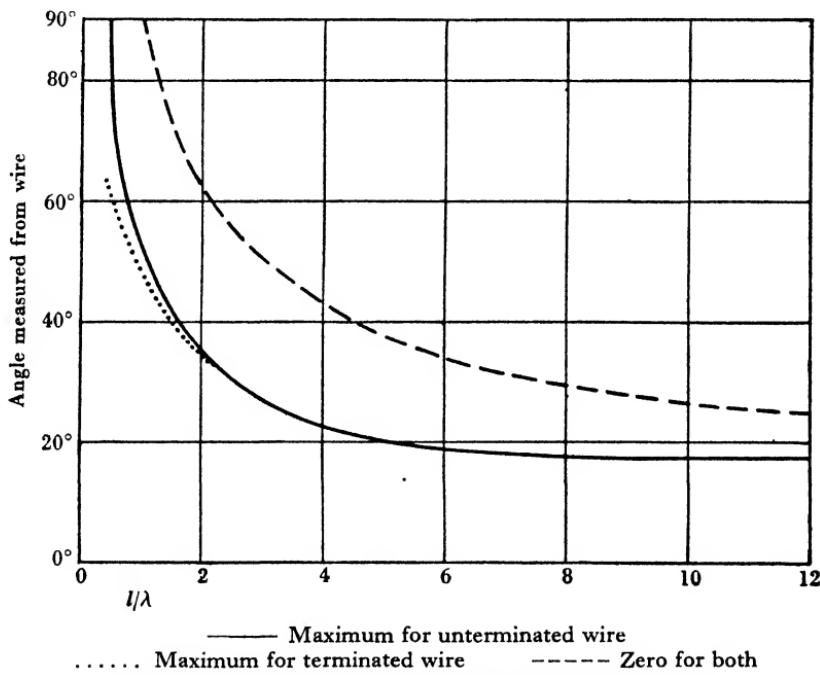
——— Maximum for unterminated wire
...... Maximum for terminated wire   – – – – Zero for both
Fig. 62. Principal maximum and first zero of wire radiator.

This differs from (5) and (6) by the factor $\sin^2\theta/(1 - \cos\theta)$ when l is an integral number of half-waves. This is a slowly varying factor having the value 2·0 at $\theta = 0$ and 1·0 at $\theta = \frac{1}{2}\pi$, so that when $l \gg \lambda$ the maxima are nearly the same as obtained from (5) and (6). The zeroes are, of course, exactly the same. For small values of l the first maximum is shown in fig. 62. The polar diagram for $l = \frac{1}{2}\lambda$ is shown in fig. 86 of Chapter 11. The effect of this factor is to shift

the maximum from the value 90° for an unterminated wire to about 65°. When $l > 2\lambda$, however, the difference in polar diagram in the region of the main lobe is small. When $l \gg \lambda$ we note that the principal maxima of the patterns given by (5), (6) or (8) occur for small values of θ. When θ is small all three reduce to the form

$$\text{const. } \sin\left(\frac{\pi l \theta^2}{2\lambda}\right)\Big/\theta. \tag{9}$$

The maxima of this expression are given by the solutions of the equation

$$\tan x = 2x, \tag{10}$$

where $x = \pi l \theta^2 / 2\lambda$.

The solution of (10) corresponding to the first maximum is $x = 1 \cdot 15$, giving

$$\theta_m^2 = 3 \cdot 3 \lambda / \pi l, \tag{11}$$

or

$$\theta_m = 1 \cdot 05 (\lambda/l)^{\frac{1}{2}}. \tag{12}$$

For most practical purposes the approximate form

$$\theta_m = (\lambda/l)^{\frac{1}{2}} \tag{13}$$

is a good enough approximation for the position of the first maximum when $l \gg \lambda$.

The first zero is given by

$$\theta = (2\lambda/l)^{\frac{1}{2}}. \tag{14}$$

The width of the main lobe between 3 db. points is therefore of the order of $(\lambda/2l)^{\frac{1}{2}}$ radians when $l \gg \lambda$.

The most striking difference between (8) and (5) or (6) is that whereas the open-ended wire has a polar diagram symmetrical

(a) Unterminated   (b) Terminated
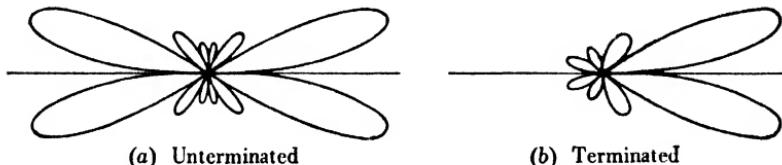
Fig. 63. Polar diagrams of wire radiators.

between θ and $(\pi - \theta)$ the terminated wire has not, radiation being much smaller for values of $\theta > \frac{1}{2}\pi$. The terminated wire produces only two main lobes in the 'forward' direction. This is illustrated in fig. 63. We note, however, that a long wire, even with an open end, will tend to behave more like (b) than (a) as the length is increased.

The theoretical front-to-back ratio in the direction of the principal maximum is approximately equal to $4/\theta_m^2$ or $4l/\lambda$, unless l is an odd multiple of $\frac{1}{4}\lambda$ when it is much greater. In practice this high value is not generally realized.

8·1·3. *Physical explanation of optimum condition*

Equation (13) has received a simple physical explanation by Bruce(3). Consider a wave incident on a wire as shown in fig. 64. Suppose we neglect reflexions at the ends. Then each element of the wire will be excited by the wave and will send a wave along the wire to A. If AC is perpendicular to the wave front and C is the projection of the free end B or AC, the difference in phase in arrival of the contributions of the elements at A and B will be $2\pi(AB-AC)/\lambda$. The vectors corresponding to the elements of the wire will form an
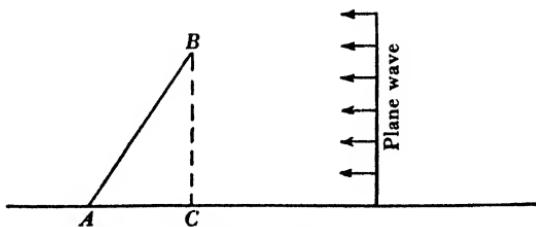
Fig. 64. Action of long-wire aerial.

arc of a circle. The resultant vector will be greatest when it is a diameter of the circle, i.e. when the contributions from A and B are 180° out of phase. Another way to look at this is to regard BC as part of an aperture plane. If in passing from B to C the phase changes by 180°, B can be regarded as lying at the edge of the first Freznel zone centred on C. It is a well-known principle in optics that maximum radiation is received from an aperture just the size of the first complete Freznel zone. The condition that the contribution from B and C should be 180° out of phase is just

$$(AB-AC)=\tfrac{1}{2}\lambda \tag{15}$$

or
$$\frac{2\pi l}{\lambda}(1-\cos\theta)=\frac{\pi}{2}, \tag{16}$$

which reduces to (13) when $l \gg \lambda$. For long-wire aerials consisting of single wires or groups of wires (15) is widely used for design purposes.

The condition (16) for maximum performance is only valid for long wires. In this case the phase difference on the left of (16) changes rapidly with θ so that variation of the magnitudes of the elementary vectors with θ may be neglected as a first approximation. As θ tends to zero the contributions from all the elements of the wire will, of course, be in phase, but each will tend to zero as the electric force will become perpendicular to the wire.

There is, however, a method of making the contributions from each of the elements add up in phase. If the velocity of propagation along the wire had a value V greater than the velocity of light in free space c, the contributions would add in phase provided

$$\cos \theta = c/V.$$

If V is sufficiently great to give a fairly large value of θ a considerable improvement in reception would be obtained because the optimum angle would be increased resulting in an increase of the $\sin \theta$ factor arising from the inclination of the electric field to the wire.

The desired increase in V may be obtained by inserting series condensers in the wire. A practical form of such an aerial has been described [4]. This consists of a pair of twisted wires, each alternate one of which is cut at intervals of $14\frac{1}{2}$ in. The capacity between the overlapping sections provides the loading required to increase the velocity of propagation along the aerial.

8·1·4. *Radiation resistance and power-gain of wire radiator*

Let us consider first a thin fairly short-wire aerial n half wavelengths in length, n being a small integer. If i_0 is the maximum value of the current we may define the radiation resistance r_0 in terms of the power P radiated by means of the relation

$$P = \tfrac{1}{2} i_0^2 / r_0. \tag{17}$$

If the aerial were fed at a current maximum so as to produce the current (3), then r_0 would just be the resistive part of the impedance as seen at the feed point. If the aerial is fed at one of its ends the input impedance to the above approximation would be infinite. The same considerations apply, however, to a half-wave aerial. As the length of the aerial is increased the resistive part of the input impedance will pass through a series of maxima and minima near which the reactance will change sign. The maxima and minima will move together and the reactance will decrease as the length is

increased, the input impedance gradually becoming independent
of length and tending to the expression (1). There is no simple
expression for the input impedance when l is not much greater
than λ. For the radiation resistance, however, as defined in (17),
we may proceed as in §4·1. The calculation is simple when l is an
integral number of half waves and gives

$$r_0 = 30[\gamma + \log_e(4\pi l/\lambda) - \text{Ci}(4\pi l/\lambda)], \tag{18}$$

which reduces to the value given by equation (31) of Chapter 4
when $l = \frac{1}{2}\lambda$.
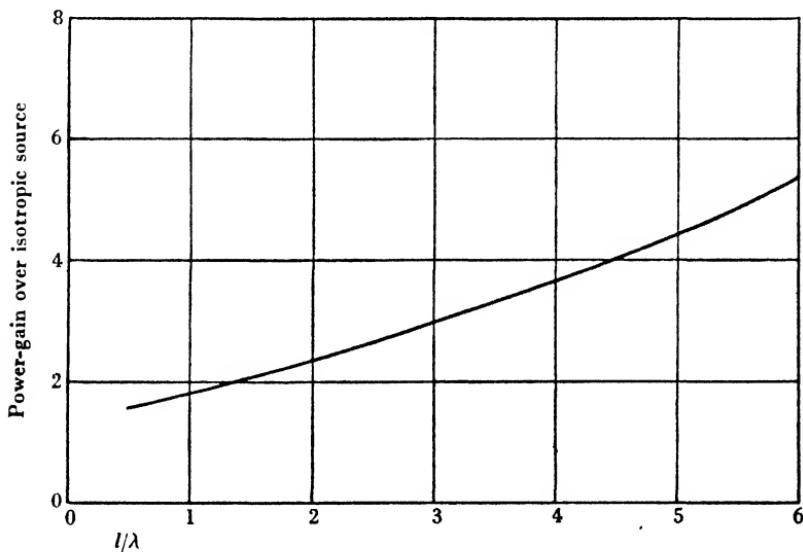
Fig. 65. Power-gain of single-wire radiator.

The power-gain of a wire radiator may be obtained from its polar
diagram by means of equation (5) of Chapter 3 by numerical
integration. This is shown in fig. 65 for a wire having various lengths
up to 6λ. For the longer lengths the value is only approximate, as
the current is no longer even approximately of form (3). When
$l \gg \lambda$ perhaps the easiest way to estimate the power-gain is to
compare the relative field strength at the maximum of the first lobe
for a given input power with that from an isotropic source. We may
readily calculate the input power now as we know the input
impedance, which is given by (1).

From (9) and (13) we have, if $l \gg \lambda$,

$$E_{\max} = \frac{120 i_0}{R} \left(\frac{l}{\lambda}\right)^{\frac{1}{2}}. \tag{19}$$

Then
$$\tfrac{1}{2} G Z_i i_0^2 = GP = \frac{E_{\max.}^2}{240\pi} \times 4\pi R^2, \tag{20}$$

giving
$$G = \frac{480}{Z_i} \frac{l}{\lambda}, \tag{21}$$

Z_i being given by (1). When l is not large compared with λ, G is obtained from (20) and (8) using an exact value of $\theta_{\max}$.

For longer values of l the power-gain increases more slowly than indicated by (21) owing to attenuation of the current wave down the wire. In practice the values found are always somewhat less than given by (21).

8·2. Combinations of wire radiators

Various combinations of wires have been used to provide aerials with more directivity than that produced by a single wire. The most

(a) V aerial   (b) Diamond aerial

Fig. 66. V and diamond aerials.
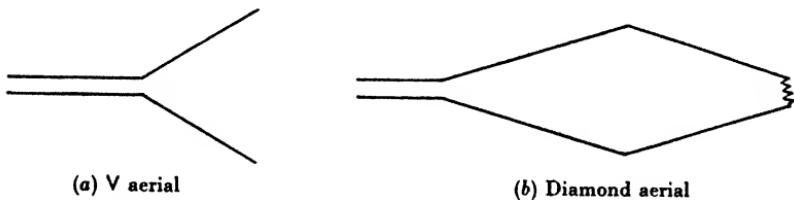
(a) V aerial   (b) Diamond aerial

Fig. 67. Typical polar diagrams of diamond and V aerials.

commonly used for horizontal polarization are the V aerial and the diamond or rhombic aerial. These are shown in fig. 66 and have been widely treated in the literature. The main lobes from all the wires making up the array are made to coincide along the desired direction. Typical polar diagrams for V and diamond aerials are shown in fig. 67.

8·2·1. *Diamond or rhombic aerials*

By combining the contributions from the four sides of a diamond aerial, with their appropriate phases, the field strength in the plane of the aerial may be found. For V aerials and diamonds with sides many wave-lengths long the angles between principal direction and the wires forming the aerial are, of course, just given by (13) or the equivalent form (15). We note that provided condition (15) is satisfied the contributions from all the wires add in phase. In the principal direction the change in phase in the contributions from the beginning and end of a wire is π, and this is just balanced by the effective change of phase π at the corner due to the bending of the wire. The full expressions for the field strength are somewhat cumbersome but have been given frequently in the literature [5]. On the axis of the diamond aerial the field strength E may be deduced at once from (8). It has, in fact (in the forward direction), four times the value given by (8), if 2θ is the angle between the terminated legs. The value in the backward direction is obtained by replacing θ by $\pi - \theta$. The polar diagram up to the first zero, in the plane at right angles to the plane of the aerial, is very similar to that in the plane of the aerial itself.

The power-gain of a diamond aerial will be approximately eight times that of a wire of length equal to that of one side of the constituent rhombus. The contributions from the four sides will add in phase, but the input impedance will be doubled. Attenuation reduces the contributions from the terminated sides, and in practice a smaller value is found, of the order of six times that given by (21).

One interesting application of a diamond aerial in radar was its use to produce a very narrow beam in anti-phase with the main transmission to suppress the fixed echoes from balloon barrages and other prominent objects.

8·2·2. *Inverted V aerials*

For vertical polarization inverted V aerials are frequently used rather than open-ended wires. The inverted V is just half a diamond having one end terminated and is shown in fig. 68. It is an extremely easy aerial to erect as it requires only one mast to support it. This mast is preferably made of insulating material. The height h of the apex of the V, if $l \gg \lambda$, is given by

$$h = (l\lambda)^{\frac{1}{2}}, \tag{22}$$

in order to satisfy (15). This only holds, however, if there is no sagging of the wire. In taking this into account one should aim at retaining the relation (15) with AB replaced by the actual length of the wire. For very short waves elevated diamonds and open-ended V's have been used for vertical polarization and have even been used as units in more complex arrays.

8·2·3. *Effect of the ground*

The free-space polar diagram of a diamond or V aerial will, of course, be modified by the presence of the ground in accordance with the principles given in Chapter 3. For the purpose of calculating the vertical polar diagram, a diamond or V aerial with axis horizontal may be taken as a point source at the height of its axis, on which the phase centre of the array lies, as may be easily seen from the symmetry of the arrangement. To estimate the gain available in practical conditions somewhat elaborate calculations are necessary unless the angle of arrival of the desired radiation is very small. In this case the gain over a half-wave dipole placed at the height of the axis is just the free-space gain over that of a half-wave dipole.

In the case of an inverted V aerial, however, it is much more difficult to estimate the effective height of the aerial, and the free-space gain is of little significance as the aerial is always used with its ends near ground-level. We should, in fact, compare it with an earthed unipole, but because of its greater effective height the gain at small angles of elevation will in general be greater than the free-space gain compared with a unipole. It may be shown by evaluating the phase factor in (8) and combining with an image in anti-phase that the effective height of the V aerial is just $\frac{1}{2}h$, as may, indeed, have been expected. The gain at small angles of elevation will then be four times that of a long wire equal to one edge of the V over a half-wave dipole placed at a height equal to half that of the apex of the inverted V.

8·3. Use of inverted V aerials with Gee

The inverted V aerial being simple to erect and having good directional properties was ideal as a receiving aerial at Gee 'slave' stations(6). The only new design points which arose were associated with the terminations of the aerial. One of the most interesting features of the travelling wave aerial is its relative insensitivity to

frequency changes. It is, in fact, an ideal broad-band aerial. When the Gee frequency bands were extended, these aerials had to work over enormous frequency ranges, and did so very satisfactorily. In order to do this, however, the aerial must be properly terminated. A difficulty arises when very short waves are used, since the earth, particularly when dry, is not a good conductor and behaves much more like a dielectric. Burying the remote end of the aerial wire is often found to be quite ineffective, and a much more satisfactory operation is obtained by the use of a dipole earth. The aerial wire is connected by means of a resistor to the centre point of a fat $\frac{1}{2}\lambda$ dipole supported $\frac{1}{2}\lambda$ or more above the ground as shown in fig. 68.

8·3·1. *The resistive termination*

For receiving purposes it is much easier to provide a resistive termination than for a transmitting aerial. A small non-reactive resistor is adequate for the former, whereas for the latter a large resistor capable of dissipating about half the transmitter power is required, and it is very difficult to make this non-reactive. In any case it is usually desirable to enclose the terminating resistor in a weatherproof box, and the capacity introduced by this and the insulator to which the aerial wire is attached needs to be removed by an inductance in shunt with the terminating resistor. Unless this is done, reflexions are likely to take place from this end. The value of the resistor required depends on the wave-length and diameter of the wire and is given by equation (1). For λ/a equal to 6000 the value is 400 ohms. These are about the values normally used in practice. At the input end the impedance of a properly terminated V aerial is about equal to this value and must be matched to a concentric line leading to the receiver. This line usually has a characteristic impedance of about 60 ohms. The match may be accomplished by means of an ordinary condenser and inductance L-section which, with a match of this order, is effective over a moderate band width. The matching unit is usually mounted in a metal box which is 'earthed' by means of a dipole earth similar to that used at the remote termination. It is desirable to keep this match good so as to avoid reflexions which would spoil the pulse shape in the case of radar reception or would introduce distortion in the reception of television signals. The dimensions of the aerial, including the $\frac{1}{2}\lambda$ earth, are made correct for the mid-frequency

of the band over which it is required to operate. The distance d is determined by the height of mast available but should, if possible, be at least 5λ. Then $(h - h') = (l\lambda)^{\frac{1}{2}}$ (fig. 68). An interesting modification to produce a two-directional inverted V aerial which is still aperiodic was introduced by R. J. Cary. The V is terminated at
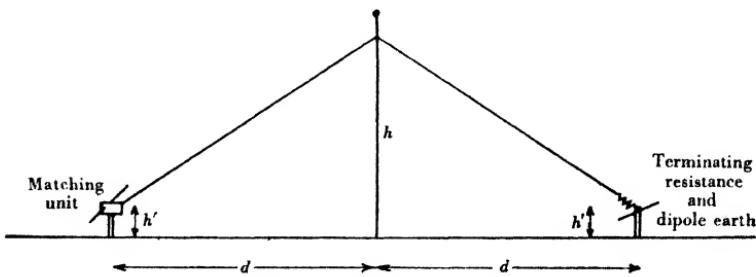
Fig. 68. Inverted V aerial.

each of its ends and a balanced feed is taken from the apex. A balance-to-unbalance transformation by means of a lattice network is used, and the signals are led down the mast by means of a concentric line. This does not appear to upset the working of the aerial, each half of the V acting as a long-wire terminated aerial.

REFERENCES

(1) Ladner, A. W. and Stoner, C. R. *Short Wave Wireless Communications*. Chapman and Hall, 4th ed. (1942).

(2) Schelkunoff, S. A. Theory of antennas of arbitrary size and shape. *Proc. Inst. Radio Engrs, N.Y.*, **30**, 493 (1942).

(3) Bruce, E. Developments in short-wave directive antennas. *Proc. Inst. Radio Engrs, N.Y.*, **19**, 1406 (1931).

(4) Cork, E. C., Pawsey, J. L. and Manifold, M. B. British Patents nos. 414, 490, 493 and 758. Also *Wireless World*, 31 August 1939.

(5) Bruce, E., Bech, A. C. and Lowry, L. R. Horizontal rhombic antennas. *Proc. Inst. Radio Engrs, N.Y.*, **23**, 24 (1935).

(6) Smith, R. A. *Radio Aids to Navigation*, Chapter 6. Cambridge University Press (1947).

Chapter 9

BROADSIDE ARRAYS FOR 1½-METRE GROUND RADAR

9·0. Introduction

The need for wave-lengths much shorter than those in the 10 m. band was very soon appreciated in the early days of radar. The lack of low-angle cover was one of the chief defects of the 10 m. radar. It will be seen from equation (20) of Chapter 2 that the first maximum of the vertical polar diagram can be lowered by either increasing the height h or decreasing the wave-length λ. It was usually impossible to increase the height of the aerials so that the development of shorter wave-lengths was necessary. Wave-lengths in a band around 1½ m. were adopted for wide use. These led not only to considerable advances in ground radar but to the use of radar in aircraft. Aerials for the latter application will be dealt with in a later chapter.

The chief advance in radar technique made with the introduction of 1½ m. wave-lengths was the use of relatively narrow beams. These made possible the use of direct displays such as the Plan Position Indicator and also led to a very much higher precision in determination of azimuth[1]. In the present chapter we shall be concerned with the design of broadside arrays to produce this beaming. An account of the experimental work carried out in the development of these aerials has been given by Taylor and Westcott[2].

9·1. Broadside arrays for transmitting and receiving

The first arrays used for C.H.L. radar[1, 3] were broadside arrays of dipoles. These consist of four bays, each bay having four full-wave dipoles, as shown in fig. 69, the dipoles being made of $\frac{5}{16}$ in. diameter copper tube and fed by vertical tubes of $\frac{3}{8}$ in. diameter spaced 1½ in. apart. The horizontal spacing between the centres of successive bays is $\frac{5}{4}\lambda$, and they are supported on a wooden framework, in front of a wire-netting reflector. In order to avoid losses the dipoles are clamped in insulating sleeves which are arranged to be near voltage minima at the centre of each half dipole. The dipoles are attached at intervals of $\frac{1}{2}\lambda$ along the vertical feeder

and, in order that they may radiate in phase, they are connected alternately to the two vertical lines, as can be seen in the figure. Full-wave aerials of this type, rather than centre-fed half-wave aerials, are used because they have a much higher impedance, and when several are connected in parallel the resulting impedance is near that of an open-wire feeder.

In the transmitting aerial each bay is fed by an open-wire transmission line, made of 200 lb./mile copper wires spaced $\frac{7}{8}$ in. apart, giving an impedance of 330 ohms. This is tapped on to the vertical feeder to the bay at a point where the resistive component is 330 ohms and the resulting susceptance is then removed by placing a shorting bar farther down the vertical feeder, to act as an inductive stub. The 330-ohm copper-wire feeder line has insulating spacers at intervals of about $2\frac{1}{2}$ in. in order to maintain the $\frac{7}{8}$ in. separation. Initially rather inferior dielectrics were used for the spacers, but later they were replaced by distrene which gave much less loss. Losses were further reduced by replacing the simple rectangular block spacer by an annular ring in which the dielectric was in a much weaker part of the field.

Fig. 69. Bay of four full-wave dipoles.

The receiving array is constructed of similar bays of stacked dipoles, but in order that concentric cable may be used for the receiver input, the matching at the array includes a conversion from balanced to unbalanced feeder. Various line circuits have been used for this transformation. The original method used was that employing a half-wave loop and quarter-wave transformer, illustrated in fig. 70 (a). This has some advantages from a mechanical point of view, but has the disadvantage that the unbalanced output impedance is equal to the balanced input so that, when 80 ohm concentric cable is used, a point on the vertical feeder of impedance 80 ohms has to be found. This is at one extreme of the range and adjustment is fairly critical. As an alternative, the simple half-wave loop, shown in fig. 70 (b), may be used. This requires careful design

to get good bonding between the cable sheaths, but has the advantage of producing a 4 : 1 impedance transformation. In this case, there-fore, 80 ohm cable matches into a 320 ohm point on the array, and this is much easier to obtain over a wide range of frequency. An alternative method of transformation suitable for this wave band is the well-known 'Pawsey stub' shown in fig. 70 (c). This has the
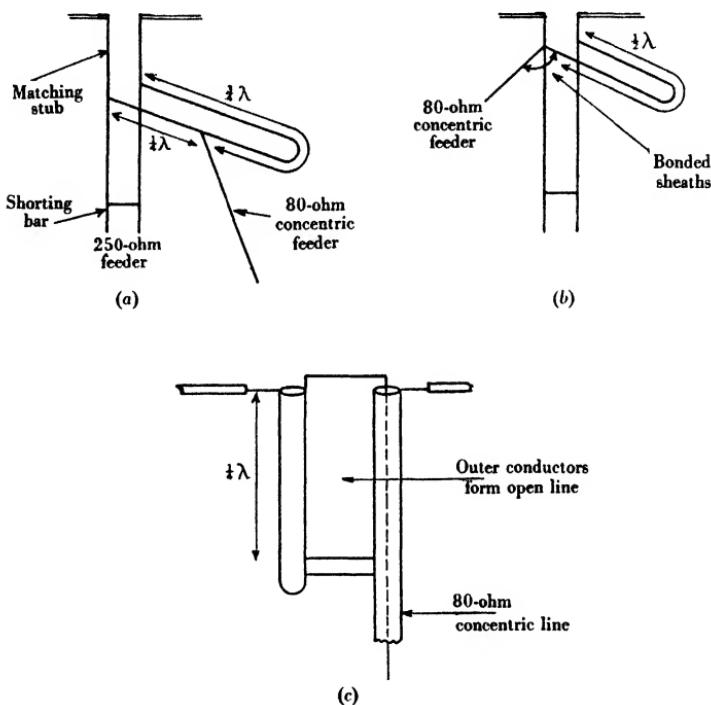
Fig. 70. Balance-to-unbalance transformer using lines.

advantage of giving a balance-to-unbalance transformation over a much wider band of frequencies but takes up more room than the simpler methods using concentric lines.

Fig. 71 (a) shows the horizontal polar diagram which is obtained with a four-bay aerial of the type described. In operation the trans-mitting and receiving aerials are rotated together until a target is illuminated by the narrow horizontal beams. In order to obtain accurate 'direction finding', use is made of the 'split-beam' principle[1]. It is arranged that the receiving array polar diagram

changes in rapid succession between two alternative beams as shown in fig. 71 (*b*). These two beams are directed respectively slightly to either side of the normal to the array, and only when the reflecting target lies on the normal to the array will the received signals from the two beams be equal. This equality can be judged very exactly by displaying the signals side by side as amplitude deflexions of

(a)

(b)
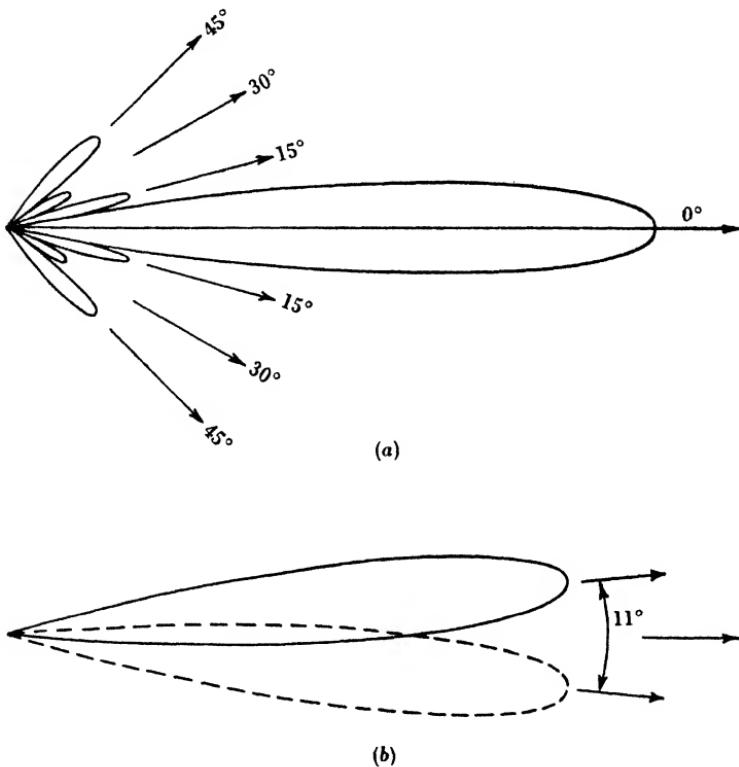
Fig. 71. Polar diagrams of 1½ m. broadside arrays.

a cathode-ray tube trace. The switching between the beams is carried out at about 20 cyc./sec. by means of a 'split motor' which produces two alternative paths from the four bays of the array to the receiver. This is shown schematically in fig. 72. The motor shaft carries a small cam which alternately connects the inner conductors of the concentric lines A, B and C, D to the main input to the receiver. With terminals A, B connected, the lengths of the feeders to the four bays are, respectively, $L-x$, L, $L+x$, $L+2x$,

giving a phase displacement across the aerial. As we have seen in § 2·4 this produces a beam slightly to the left of the normal; with C, D connected, the feeder lengths are as above, but in the reverse order. The lengths x, $3x$ of the loops are chosen to give a split angle of about 11°. It will be realized that use of the split system introduces a slight loss of gain, since the target is never in the maximum of the lobe, of both transmitting and receiving arrays, but the additional accuracy and ease of 'following' more than compensates for this.
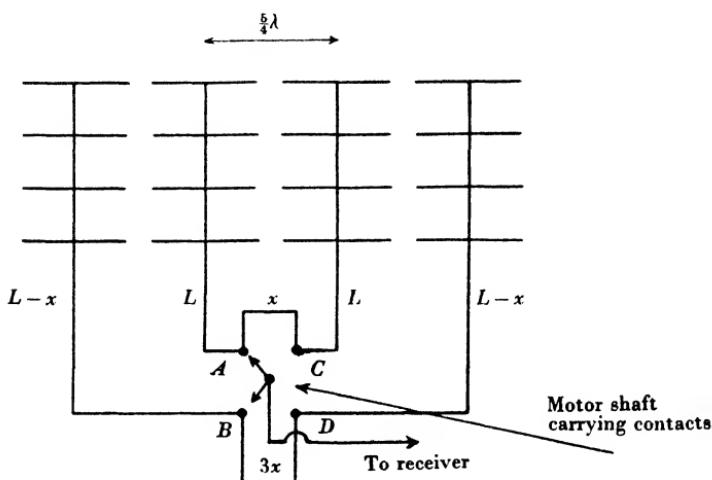
Fig. 72. Arrangement for producing split beam.

9·1·1. *Power-gain of the broadside array*

As we have seen in Chapter 2, to a first approximation, neglecting mutual impedances, the power-gain of a broadside array such as described above relative to a $\frac{1}{2}\lambda$ dipole is equal to N, the number of half-wave dipoles, or $2N$ when a reflecting screen is used. This would give a value of 64 for an array as described, in comparison with the figure of about 70 obtained from experiments.

A better estimate may be made in terms of the area of the broadside array, using equation (13) of Chapter 2, giving the result that a gain of about $8A/\lambda^2$ (relative to a half-wave dipole) is obtainable by filling an area A with half-wave dipoles, placed about $\frac{1}{2}\lambda$ apart, and backing with a reflecting screen. Since $A = 9\cdot5\lambda^2$, allowing $\frac{1}{4}\lambda$ overlap at top and bottom, this gives a value of 76 for the gain over a $\frac{1}{2}\lambda$ dipole.

9·2. Reflector screens

With the arrays described in §9·1 it is necessary that radiation should be sent out in the forward direction only, otherwise not only will the coverage be obscured by echoes from hills and buildings, but also ambiguities in direction finding will result. Hence, as mentioned, a reflecting screen is used. At a wave-length of 1½ m. one of the most convenient forms of reflecting screen is a vertical screen of ordinary wire netting, preferably of mesh as small as ½ in., erected behind the dipole array. The twists in the wire netting should be parallel to the dipoles, i.e. horizontal.

A screen of this type is sometimes spaced $\frac{3}{8}\lambda$ behind the dipoles, this spacing having been shown by theory, and by experiments with a single bay of dipoles, to give maximum forward gain (see Chapter 5). However, for a four-bay array, the forward gain is almost equally good with a reflector spacing of $\frac{1}{8}\lambda$. This reduced spacing has other factors to commend it, in particular, the leakage round the edges of the screen is about five times as small, giving an improved back-to-front ratio. Also, the small spacing is much more convenient mechanically. Instead of building up a double frame structure to carry reflector and dipoles respectively, it is possible to support the dipoles by insulators carried on a single frame which bears the reflecting screen.

9·3. Development of common aerial systems

Probably the main disadvantage of the early C.H.L. stations was the use of two aerial systems which had to be rotated in synchronism. The aerials were mainly hand-turned, and as a result searching was monotonous and exhausting and errors in 'following' reduced the performance of the system.

Various methods of overcoming the above difficulties were tried: first of all two broadside aerials on one gantry were tried but offered too much windage; next, a Yagi aerial for transmitting was mounted on top of the receiving broadside array. The latter system was successful to a limited extent, but the performance was inferior owing to the gain of the Yagi aerial being less than that of the broadside array and to the fact that the different heights of the two made the vertical polar diagrams different. Finally, a method was successfully developed, largely by C. J. Banwell[4] and R. J. Lees,

of using the same aerial for transmitting and receiving, and this procedure is now almost universally followed with all rotating aerial systems. For $1\frac{1}{2}$ m. the scheme consists of connecting both transmitter and receiver to the one broadside array and adding two 'spark-gap units' as shown in fig. 73. When the transmitter operates, both the spark gaps break down and are practically equivalent to short circuits, so that, in virtue of the properties of a $\frac{1}{4}\lambda$ transformer, they present a low admittance at A, B and the transmitter power passes direct to the aerial. In particular, the receiver gap presents a low admittance at B and prevents power from reaching the receiver.
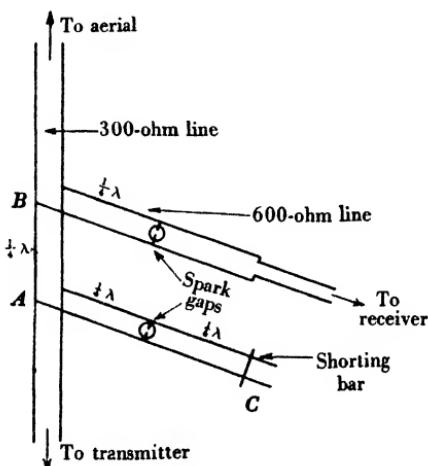
Fig. 73. Diagram of 'common aerial' unit.

Almost immediately the pulse ceases the sparks in the gaps are extinguished, and the reflected pulses can pass to the receiver. The transmitter gap presents zero admittance at A during transmission, but on reception allows the short circuit at C to be effective, producing a short circuit at A and consequently zero admittance at B. Hence, the received signal passes entirely to the receiver, and the receiving path is not mismatched by the reactance presented by the transmitter.

The requirements of a spark gap for common T and R use are low striking voltage, rapid striking, long life, low loss both when firing and quiescent and rapid deionization. The latter condition is particularly necessary when short-range echoes, received soon after the transmitter pulse has finished, have to be examined. Various

gas fillings have been used such as air, water vapour, nitrogen, hydrogen, neon and argon. The most satisfactory is argon which combines low striking voltage, long life and rapid deionization. In one of the first argon-filled gaps, made at the Cavendish Laboratory, Cambridge, a small drop of mercury was accidentally included. This improved the performance considerably, giving decreased deionization time and preventing the formation of a tungsten film on the glass, which led to adsorption of the gas. It was therefore adopted as a permanent feature. Details of the spark gaps later developed are given in Banwell's paper.

9·4. Improved feeder systems

Measurements on the original C.H.L. feeder systems showed that appreciable loss of energy occurred in the insulating spacers used for maintaining the $\frac{7}{8}$ in. separation between the two wires of the twin feeder. The loss is greatest in wet weather when water accumulates on the spacers. A combination of two methods is used to overcome this trouble. The first is the use, for as much of the feeder system as possible, of strained wire feeders which require the use of only a very few spacers. By placing pairs of shorting bars at a quarter of a wave beyond any junction point the strained wires may be attached directly to a metal array frame, without the use of any intervening insulating material. In a similar way, a long vertical feeder run may be restrained laterally by fixing double quarter-wave stubs as supports. Not only are losses much reduced, but the straight runs of feeder also give much more reproducible impedances than the meandering lengths of feeder with spacers at $2\frac{1}{2}$ in. intervals.

Where spacers are still necessary an annular ring type of distrene spacer, already mentioned in §9·1, is used. This places much less dielectric in the strong part of the field and also has a much longer leakage path, so that no trouble is experienced when the spacer is wet.

9·5. G.C.I. aerials and capacity switch

The application of $1\frac{1}{2}$ m. waves to ground-controlled interception or 'G.C.I.'[3] led to the need for aerials which would cover higher angles of elevation than the C.H.L. aerials, and which would also measure height. The C.H.L. aerial is taken as the basis and the aerial is mounted with its centre 10 ft. above the ground and

divided into two halves, 'upper' and 'lower', with their centres at $7\frac{1}{2}$ and $12\frac{1}{2}$ ft. respectively. These two halves may be fed by the transmitter either in phase or in anti-phase, and the received signals from the two halves may be examined separately. By comparing the signals from the two halves a measure of the angle of elevation and, consequently, height may be determined[2].

9·5·1. *Rotating couplings*

When the Plan Position Indicator had been developed it was possible, if the aerials could be rotated continuously, to have a continuous 'map display' of all the aircraft within the station's coverage[2]. Electrical turning gear was used, and rotating couplings without slip rings were used in order to feed the signal from the transmitter to the rotating aerials and pass the returned signals to the receiver. The most generally useful type was one using balanced transmission line, and suitable for carrying high power. This is shown in fig. 74 and consists essentially of two loops between which is inductive coupling. The coupling loops are not closed, but each is fitted at the gap with a parallel plate-matching transformer, which matches it to the 330-ohm feeder line. The matching is adjustable by varying the distance of separation between the loops. Satisfactory operation, with only 2 % power loss, may be obtained over a band of about 12 % in frequency. It will be seen from the figure that an earthed screen is fixed midway between the rings. This is introduced to make the coupling purely electromagnetic, and to remove any electrostatic effect. The latter would cause a varying impedance and signal transfer on rotation, since the electric field is uneven owing to the fact that the points on the rings opposite the gaps are at zero voltage. The screen is cut radially to reduce to a minimum the eddy currents produced by the magnetic field.

9·5·2. *The capacity switch*

As operational requirements developed it became necessary to give even better vertical cover which required the addition of further aerials. Moreover, it was no longer sufficient to feed the transmitter power into one or other of these aerials by means of a hand switch. It was necessary to use effectively three aerials at the same time, retaining the full gain of each. In order to do this it is necessary to switch both transmitter power and received signals at

high speed. This is achieved by using a high-speed, rotating 'capacity switch'.

The final G.C.I. scheme(2) uses a four-way capacity switch which feeds transmitter pulses in turn to:

(i) a four-bay four-stack top aerial with its centre at a height of 25 ft.,
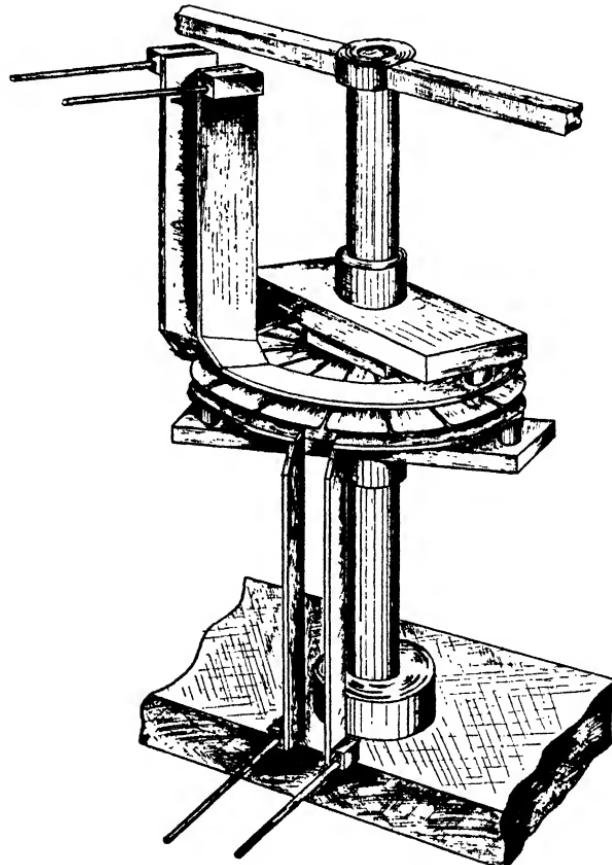
Fig. 74. Aerial coupling unit.

(ii) a four-bay two-stack middle aerial with its centre at a height of 12½ ft.,

(iii) to the middle aerial and a bottom aerial in phase, and

(iv) to the bottom aerial only, the latter being a four-bay two-stack array with its centre at a height of 7½ ft.

It is obviously very desirable that the same switch should be used

for transmission and reception, as it avoids the use of any subsidiary diode switches for the receiver. If this is to be done the switch must make only momentary contact in its four positions during the transmitter pulses, but must retain the connexions for the duration of the receiver trace. To obtain a working range of 100 miles at a transmitter recurrence frequency of 500 per sec. the change-over period
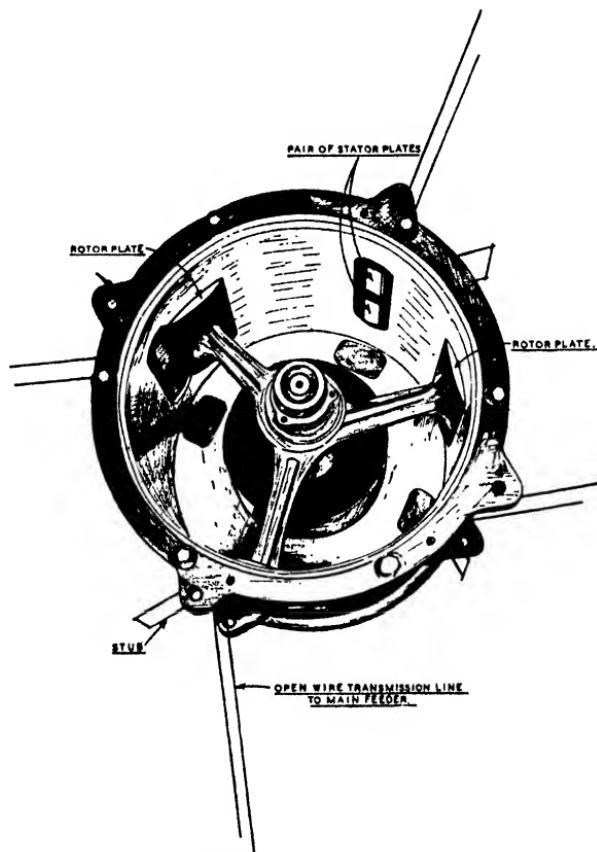
Fig. 75. General view of capacity switch.

of the switch can be only 86/186 of the pulse interval, and this factor largely determines the design(2). Essentially the switch consists of a three-pole rotor revolving within four sets of fixed plates arranged at 90° intervals as shown in fig. 75. The switching is produced by the changes in capacity, as the rotor revolves, of the condensers formed between the rotor and the stator plates.

9·6. Five-bay broadside with tapered feed

The broadside array with four bays spaced at $\frac{5}{4}\lambda$ gives rather large side lobes at $47°$, 22 % of the main lobe in field strength. In order to reduce these the bay spacing may be reduced to $1·10\lambda$ and a fifth bay introduced. A further reduction in all the side lobes, including those at $15°$ and $30°$, is obtained by using a tapered feed. A power distribution $1 : 3 : 4\frac{1}{2}$, from side to centre, is satisfactory and gives an appreciable reduction in the side lobes.

9·7. Beam swinging arrays—V.E.B.

The G.C.I. method of height finding which has been described requires an appreciable amount of level ground, free from obstacles, around the aerial in order to obtain accurate results. Uneven ground produces errors which must be allowed for by calibration, and they are, of course, different on different azimuths. Cliff-edge sites are not suitable, thus making the method inapplicable to C.H.L. stations, and low sites near the sea are unsuitable since the height calibration varies with the state of the tide. In order to overcome these difficulties a method of height finding was developed in which an aerial produces a beam which is narrow in elevation and which can be oscillated continuously in elevation rather like a scanning searchlight. This method was finally very successfully developed on centimetre wave-lengths, but an early application of it by C. J. Banwell and G. E. Bacon on 1½ m. is interesting from an aerial point of view (5).

The simplest form of the method used on 1½ m. employs a vertical stack of thirty-six centre-fed half-wave dipoles mounted in groups of four on the side of a 120 ft. tower. Centre-fed $\frac{1}{2}\lambda$ dipoles are used in this case, in contrast to the full-wave dipoles of the C.H.L. array, in order to obtain as broad a horizontal beam as possible. The impedance must be stepped up by transformers in order to obtain a reasonably high impedance for a group. Each group is mounted on a frame which carries a wire-netting reflector spaced $\frac{1}{8}\lambda$ from the dipoles which are spaced $\frac{1}{2}\lambda$ vertically. Beam swinging is obtained by phase shifting (Chapter 2). A concentric feeder is taken from each group to a phase shifter which is shown schematically in fig. 76. It consists of four circular arcs of feeder built in the form of a three-sided square trough with an insulated conductor passing down the

middle. A moving arm carrying four contacts rotated over the assembly, and therefore produced a continuous and uniform variation of phase down the nine dipole groups of the array. In this way the narrow vertical beam of width about 3° may be swung from an elevation of 0–15°. A more complicated system, designed to give
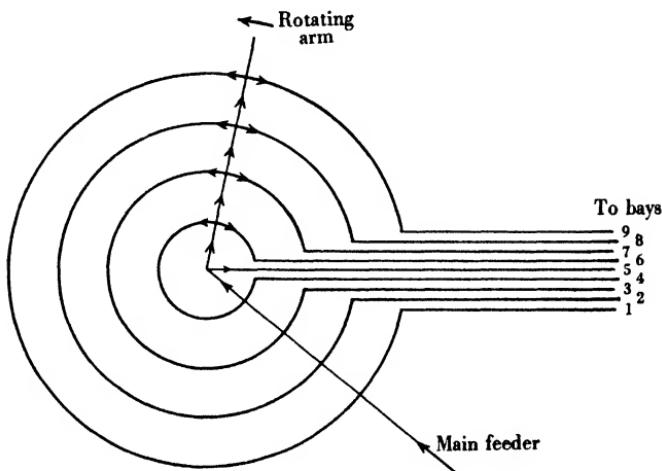
Fig. 76. Phase shifting in V.E.B. system.

a beam width of $1\frac{1}{2}°$ and to cover the same range of elevation, was later designed(5). Because of the mechanical complexity of these designs they were not used to any extent operationally till the application of microwave technique enabled use to be made of reflectors of reasonable size which could be tilted to produce the scanning in elevation.

REFERENCES

(1) Ratcliffe, J. A. Aerials for radar equipment. *J. Instn Elect. Engrs*, **93,** IIIA, 22 (1946).

(2) Taylor, D. and Westcott, C. H. Divided broadside aerials with applications to 200 Mc./s. ground radiolocation systems. *J. Instn Elect. Engrs*, **93,** IIIA, 588 (1946).

(3) Watson-Watt, R. A. The evolution of radar. *J. Instn Elect. Engrs*, **93,** IIIA, 11 (1946).

(4) Banwell, C. J. The use of common aerial for radar transmission and reception at 200 Mc./s. *J. Instn Elect. Engrs*, **93,** IIIA, 545 (1946).

(5) Bacon, G. E. Variable—elevation—beam aerial systems for $1\frac{1}{4}$ metres. *J. Instn Elect. Engrs*, **93,** IIIA, 593 (1946).

Chapter 10

YAGI AERIALS

10·0. Introduction

Yagi aerials are named after a Japanese engineer who first made use of the fact that further forward gain is obtained if more than one director is placed in front of a driven element[1]. The Yagi aerial consists of a driven element and a number of parallel directors, which may be as high as 20, with their centres in a line. A single reflector is also frequently used behind the driven element. In contrast to the directors, more than one reflector gives no appreciable increase in gain. The general arrangement is shown in fig. 77.
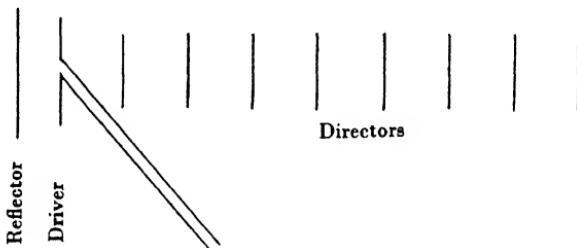
Fig. 77. Yagi aerial.

These aerials are widely used for directive arrays at wave-lengths less than 3 m., particularly for mounting on aircraft. They are much more compact than broadside arrays, but suffer from larger side-lobes. Where low windage and light weight are essential, however, they are most useful. A number of airborne applications will be described in Chapter 11. Yagi aerials are used to some extent for 1½-m. ground radar, particularly for light, transportable early-warning equipment. They have also been used to lay down beams for instrument landing of aircraft. Arrays of Yagi aerials were used to direct searchlight beams on to aircraft (equipment known as 'Elsie').

10·1. Theoretical treatment of short Yagi aerials

The theoretical treatment of such aerials follows closely on the lines given for a driver with a single parasite in Chapter 5. The problem

increases considerably in complexity as the number of elements increases, but solutions have been given by W. Walkinshaw[2] for aerials having up to four parasitic elements used in various combinations. Walkinshaw has studied the variation of power-gain, input resistance and polar diagram with variation of the reactance and spacing of the parasites. Normally only two independent values are taken so as to simplify the calculation. For example, if one parasite is used as a reflector and its reactance is varied the reactances of all the directors are assumed equal and varied. In the treatment of a driver and two parasites the reactance of each is varied independently.

The general theoretical conclusion is that the addition of further elements having negative reactance in front of the driver increases the forward gain. The addition of a single element with positive reactance behind the driver also increases the forward gain, but further addition of such elements has little effect. Optimum values are different for different spacings between the elements and also depend on the number and combination of elements used. This is illustrated in fig. 78 which shows the power-gain of Yagi aerials with two and four equal parasitic elements spaced $0 \cdot 15 \lambda$ apart.

10·2. Approximate theoretical treatment of long Yagi aerials

When more than four parasitic elements are used the calculations become too cumbersome for present methods even using an automatic calculating machine for solving the linear equations involved, and recourse has to be had to approximate methods. One such attack has been made by D. G. Reid[3] by assuming the aerial to be made up of a great many closely spaced elements leading to a continuous distribution of phase along the aerial. The optimum distribution of phase is determined for a given length of aerial, and the Yagi aerial is compared with an end-fire array of driven elements. Reid gives an approximate formula for the maximum power-gain of a long Yagi array of length l in the form

$$G = 9 \cdot 2 l / \lambda. \tag{1}$$

(Reid gives his gain relative to a Hertzian doublet; (1) refers to an isotropic source.)

The optimum current distribution for an end-fire array has been considered by Hansen and Woodward[4], who have shown that for

a uniform distribution of current the optimum conditions are obtained when the 'phase velocity' of the current distribution down the aerial is slowed up so that one more half-wave-length is obtained over the length of the aerial than would have been if the phase velocity had been c. If we assume the radiation to be propagated down the aerial with the same phase velocity as the current distribution, this is in accordance with the determination of optimum
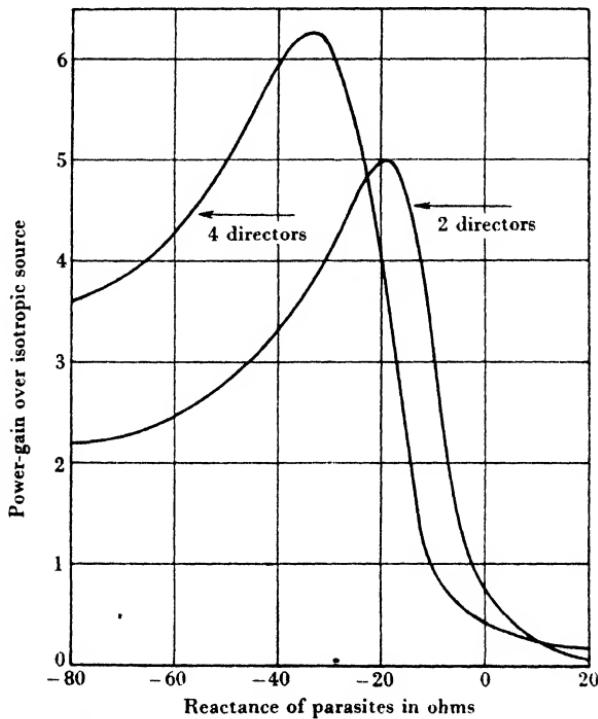
Fig. 78. Variation of power-gain of Yagi aerials
with reactance of directors.

conditions for a long-wire aerial given in §8·1·3. The maximum gain for the end-fire array over an isotropic source under these conditions is found to be $7 \cdot 99 l/\lambda$. In a Yagi aerial, however, the current distribution will not be strictly uniform. In practice a gain of only about $5l/\lambda$ is found for long Yagi aerials.

10·3. Experimental investigations of Yagi aerials

A great deal of experimental work has been carried out in order to determine the optimum dimensions of long Yagi aerials. The

number of variables is, however, so large that, even so, only very incomplete information is available. Moreover, the conditions for maximum gain, narrowest main lobe, and minimum side-lobes are not the same, and the optimum condition will depend on which of these desirable features is the most important. The most extensive experimental data available is that of Fishenden, Ryle and Wiblin (5), and most of the data given in this chapter are taken from their work.*

One of the most important facts which came to light in the early work on Yagi aerials was that unless the elements can be supported on good insulators and all poor conducting material, such as wet wood, kept from their vicinity, very poor results were obtained. Since this is difficult it is found to be much better to support all the elements at their centre on a metal rod. The driven element consists
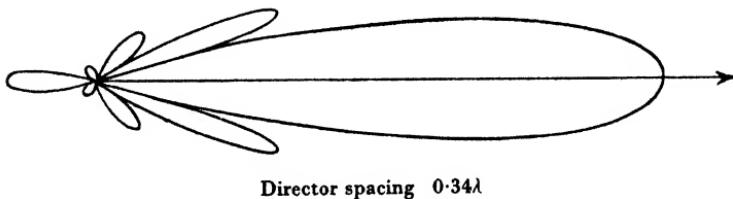
Director spacing   0·34λ
Overall length   10·5λ
Director length   0·40λ
Reflector spacing 0·30λ

Fig. 79. Polar diagram in plane of elements
of 30-director Yagi aerial.

generally of a folded dipole supported at the centre of its complete side. With such an arrangement the gain and polar diagram of the Yagi aerial is almost identical with that obtained with well-insulated elements, provided that, in estimating the length of the elements, allowance is made for the supporting rod. This simply amounts to shortening them by an amount equal to the chord of the supporting rod parallel to their length. An example of the polar diagram obtained from such a multi-element Yagi aerial is shown in fig. 79.

In the experimental data given below the directors were all of equal length and spaced 0·34λ apart. A reflector was spaced 0·25λ behind the driven element. Undoubtedly further experimental work using different spacings will show an improvement in the gain per unit length, particularly for long aerials. The value $5l/\lambda$ found

* The author is indebted for advance information prior to publication.

for long Yagi aerials is considerably less than the optimum value predicted by Reid, even though this is obtained with rather drastic assumptions. It is felt that the value $5l/\lambda$ could be improved, perhaps by using a director spacing which varies along the length of the aerial.

10·3·1. *Beam width and gain*

The most important data for an aerial will usually be the beam width and gain. Table 10 shows the variation of gain and beam width with number of directors. The figures for beam width given are for the full width at half-field strength (− 6 db.) and are obtained with a side-lobe ratio of about 30 %. Slightly narrower beams may be obtained at the expense of larger side-lobes and lower gain. The polar diagrams in the plane of the elements and that at right angles are, of course, not independently variable. The beam width given is for the plane of the elements, but that in the plane at right angles is only slightly greater.

Table 10. *Gain and beam width of Yagi aerials*

No. of directors	Beam width	Power-gain over isotropic source	Gain/unit length in λ
30	22°	—	—
20	26°	34	4·9
13	31°	25	5·4
9	37°	21	6·25
4	46°	15	9·5

The gain per unit length decreases fairly rapidly as the length of the aerial increases, and in general it is not convenient or efficient to make the aerial more than 6λ long.

10·3·2. *Director length*

Having chosen the number of directors in the aerial, table 11 gives the director length to give a reasonable side-lobe ratio of about 30 %. The figures are for a ratio of wave-length to director diameter of 160, with the directors supported near their centre by a low-capacity insulator. If the ratio is increased (smaller diameter director) the director length should be increased by about 2 % for 100 % change in ratio. If streamlined tubing is used the effective

diameter is approximately the maximum cross-sectional dimension. If a metal tube support is used the overall length is that given by table 11.

Table 11. *Optimum director length for Yagi aerial*

No. of directors	Director length (λ)
42	0·385
30	0·40
20	0·407
13	0·414
10	0·42
7	0·423
5	0·434 ·

10·3·3. *Front-to-back ratio*

At first sight it might seem that the simplest way to obtain a good front-to-back ratio for a Yagi aerial is to adjust the reflector, but, in fact, this is not so. A reflector exactly half a wave-length long is normally used, spaced $\frac{1}{4}\lambda$ behind the driven element. This spacing may be reduced to about $\frac{1}{8}\lambda$ without appreciably affecting the gain and polar diagram of the aerial, but this varies considerably the input impedance. This variation may, however, be used for matching purposes. The input reactance may be varied by changing the length of the driven element. The reflector seems to be primarily effective in improving the current distribution along the aerial, and thus improving the performance.

The front-to-back ratio may be improved by using a reflector screen behind the driven element, but this will not be effective unless the screen is large enough to subtend an appreciable angle at the last director. The use of a screen is, however, very undesirable from a mechanical point of view, as it removes most of the outstanding advantages of a Yagi aerial, namely, its compactness, light weight and low windage.

10·3·4. *Impedance of Yagi aerials*

The input impedance of a Yagi aerial is measured at the centre of the driven element. In the data given below the length of this element, a simple dipole, has been adjusted to make the input impedance resistive. Table 12 gives some typical values of input resistance. The values are rather low for matching to normal feeders,

Table 12. *Input impedance of Yagi aerial*

Reflector spacing (λ)	Director length	
	0·406λ (ohms)	0·42λ (ohms)
0·25	62	50
0·18	50	43
0·15	32	27
0·13	22	—
0·10	12	—

but may be increased by using a folded dipole as driven element. The input impedance Z then depends on the ratio of the tubing diameters for the fed and folded parts, and is four times the normal impedance when the diameters are equal (see § 1·3). The dipole length for the impedance to be resistive depends on the spacing between the two parts, and on the tubing diameters, but is of the order of 0·42λ. The factors for different diameter ratios which give the input impedance when it has been obtained with a simple dipole feed are shown in table 13.

Table 13. *Data for folded dipole*

Diameter ratio, folded : fed part	4	3	2	1	½	⅓	¼
Multiply Z by	7	6	5·2	4	3·2	2·7	2·5

10·3·5. *Band width of Yagi aerials*

Very little information is available on the band width over which Yagi aerials will operate without a serious drop in performance. One thing, however, is quite clear. The rate of fall-off in performance when the directors are too long is much more rapid than when they are too short. If a band of frequencies is to be covered it is then advisable to make the directors *less* than their optimum value. With this procedure a fifteen-element Yagi aerial may be used over a band of about 5 % of frequency without its power-gain dropping by more than 10 %. At the extreme ends of the band, however, side-lobes are likely to be considerably increased. Moreover, the input impedance is likely to change considerably, and the resulting mismatch may be intolerable, so that the permissible band width may be as low as 2 %.

10·4. Arrays of Yagi aerials

We have already noted that it is inconvenient to use Yagi aerials with length greater than 6λ, so that a gain of about 30 is all that is readily available. When more gain is required arrays of Yagi elements must be built. Two Yagi elements in the same plane do not interact appreciably if their centre lines are separated by at least $1\cdot5\lambda$. If the planes of the elements are parallel they interact to a considerable extent unless the separation has at least twice this value.

The polar diagram available may be considerably improved by the use of an array of Yagi elements. For example, if a reduction of side-lobes is the chief consideration, this may be effected to some extent by arranging that a minimum of the array factor coincides with the position of the first side-lobe of the individual elements. To decrease the width of the main beam the first subsidiary maximum of the array factor may be made to coincide with the first minimum of the elements. The first minimum of the array factor will then lie at a smaller angle with the principal direction, and will help to cut down the width of the main beam.

10·5. General principles of operation of long aerials

Although the actual detailed operation of Yagi aerials is complex their general method of operation may be seen qualitatively as
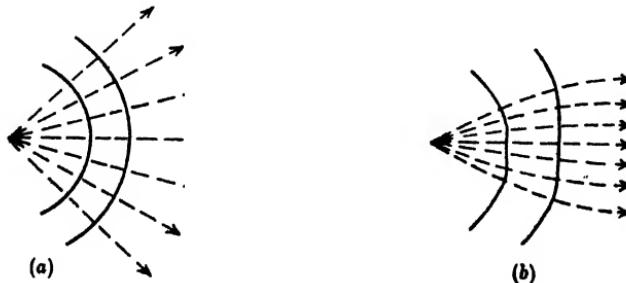
Fig. 80. Effect of reduction of phase velocity
along axis of aerial.

follows. The effect of the directors may be represented as a slowing down of the phase velocity of propagation along the length of the aerial. For an isotropic source the constant-phase surfaces will be as shown in fig. 80(a) and the flow of energy will be uniformly distributed in space. The effect of reducing the velocity of propaga-

tion along the axis is shown in fig. 80(*b*). As a result more energy is concentrated towards the axis of the aerial.

Long-wire aerials such as the diamond or inverted V may be regarded as devices for slowing down the phase velocity along their axis. As the waves are guided along the wires with velocity *c* the effective velocity along the axis is $c \cos \theta$, where θ is the angle the wires make with the axis.

10·6. Dielectric rod aerial

These considerations would appear to indicate that a rod of dielectric might be used to decrease the phase velocity along its
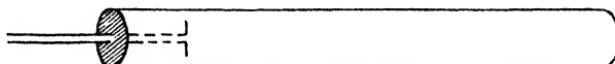
Fig. 81. Dielectric rod aerial.

axis and so provide a directive aerial as shown in fig. 81. In fact, this is so, and although such aerials have not been much used in this country they have been extensively used in Germany(6). Their general behaviour is very similar to that of Yagi aerials. They may be built into arrays to form the so-called 'Polyrod' aerials.

REFERENCES

(1) Yagi, H. Beam transmission of ultra-short waves. *Proc. Inst. Radio Engrs*, *N.Y.*, **16**, 715 (1928).
(2) Walkinshaw, W. Theoretical treatment of short Yagi aerials. *J. Instn Elect. Engrs*, **93**, IIIA, 598 (1946).
(3) Reid, D. G. The gain of an idealised Yagi array. *J. Instn Elect. Engrs*, **93**, IIIA, 564 (1946).
(4) Hansen, W. W. and Woodward, J. R. A new principle in directional antenna design. *Proc. Inst. Radio Engrs*, *N.Y.*, **26**, 333 (1938).
(5) Eishenden, R. M., Ryle, M. and Wiblin, E. R. Design of Yagi aerials. *J. Instn Elect. Engrs* (in press).
(6) Mueller, G. E. and Tyrrell, W. A. Polyrod antennas. *Bell System Tech. Jour.* **26**, 837 (1947).

Chapter 11

AIRCRAFT AERIALS

11·0. Introduction

The introduction of very short waves for communications between ground and aircraft and for radar led to a considerable development of aerials for use in aircraft. For wave-lengths less than a few metres the presence of the aircraft framework has an important influence on the performance of aerials mounted on the aircraft. If the skin of the aircraft is metallic, the wings and fuselage act as screens, and although there will be some diffraction, this screening will have a marked effect on the aerial polar diagram. For example, direction finding from an aircraft is impossible at wave-lengths in the metre band due to interaction between the aerials and aircraft structure, and only homing is possible by using the symmetry properties of the aircraft.

Aircraft aerials may be divided roughly into three classes:

(1) Homing aerials such as are used for various radar applications.

(2) Aerials producing a narrow beam in a fixed direction, usually in the line of flight or at right angles to it.

(3) All-round aerials producing as uniform a polar diagram as possible in the horizontal plane. Such aerials are required for communications or for navigational aids.

We shall consider these in turn.

At wave-lengths of the order of $1\frac{1}{2}$ m., aerials consisting of simple dipoles and reflectors offer very little drag to the aircraft, except at very high speeds (say over 450 m.p.h.). With the increasing speeds of modern aircraft the use of any form of external aerial is becoming more and more undesirable, and the trend is towards the use of aerials which do not project beyond the lines of the aircraft structure. We shall consider such aerials, known as 'suppressed' aerials, later in the chapter.

Another problem in the design of aircraft aerials is to find a form of aerial which may be attached to any type of aircraft and still maintain its expected performance. Although some progress towards this ideal has been made, for many purposes it is necessary to find

by experiment the best positions for the aerials for each kind of aircraft. This procedure involves a large amount of labour.

11·1. Homing aerials

The use of homing aerials for radio navigational purposes has been described in some detail in another book of this series[1]. Similar aerials have also been widely used for $1\frac{1}{2}$ m. airborne radar such as A.I.* and A.S.V.* The design and positioning of such aerials have been fully described in an extensive paper by B. Russell[2].

Although the taking of accurate bearings from an aircraft is impossible in the metre band, it is possible, by making use of the symmetry properties of the aircraft, to tell whether a 'target' is to
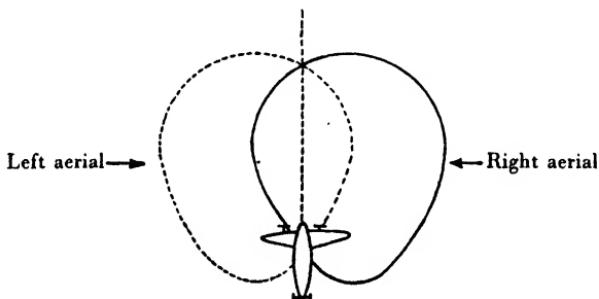
Fig. 82. Homing by use of overlapping polar diagrams.

right or left or dead ahead. This is done by using overlapping polar diagrams from two aerials, one to the right and one to the left of the aircraft, the signals from the two aerials being equal when the 'target' is dead ahead. By a knowledge of the shape of the polar diagrams a rough estimate of the extent to which the 'target' is to right or left may be obtained. This procedure is illustrated in fig. 82.

By using the wings of an aircraft as screens similar overlapping polar diagrams may be obtained in the vertical plane and homing in elevation may thus be achieved. The method is illustrated in fig. 83.

For homing in azimuth the accuracy available depends on how much coverage in the horizontal plane is required. If aerials producing narrow beams in the forward direction are permissible, considerable accuracy may be achieved, as is done, for example, in

* A.I., air interception. A.S.V., detection of surface vessels from the air.

A.S.V. radar. For navigational purposes, however, 'all-round' cover is required and no substantial beaming is permissible. In this case a homing accuracy of only ± 5° is available, but this is in general adequate.

(a)

(b)
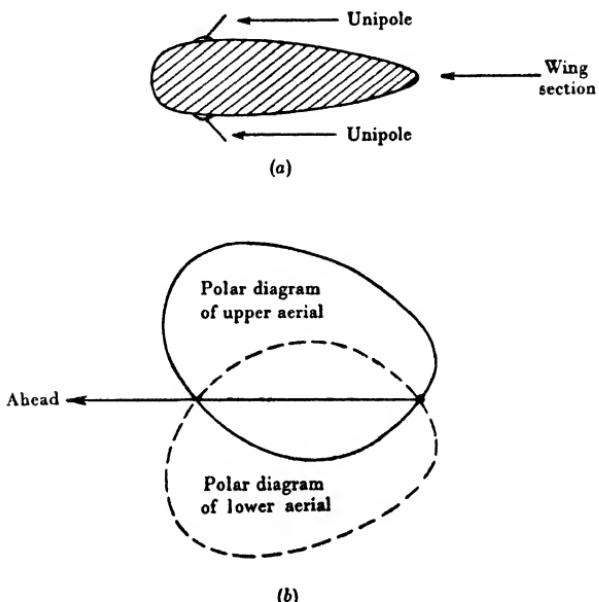
Fig. 83. Use of wing section to obtain homing in elevation.

11·1·1. *Homing aerials having wide coverage*

There are two methods in which the overlapping polar diagrams as shown in fig. 82 may be obtained. Use may be made of the aircraft structure acting as a screen to obtain the diagrams, or an aerial array itself producing the necessary pattern may be used. The second method is preferable as it is more readily adapted for use on many types of aircraft, whereas the former requires each type to be considered as a separate problem. In figs. 84 and 85 are shown examples of the first method for vertical and horizontal polarization respectively. In order to obtain successful operation by the methods shown it is necessary for the fuselage to extend for more than a wave-length in front of the propellers. For horizontal polarization the $\frac{1}{2}\lambda$ unipoles are swept back at an angle of about 45° to the line of flight. The director shown in fig. 84 is used to reduce the distortion due to the propeller and engine nacelle, and to increase
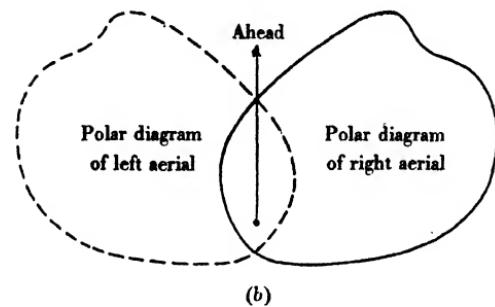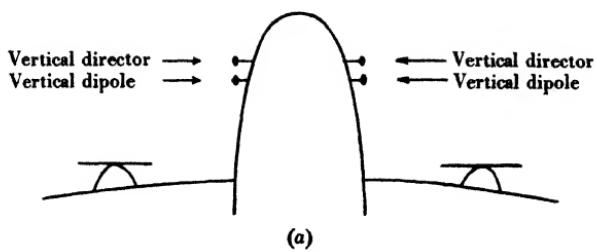
Fig. 84. Use of fuselage to obtain overlapping polar
diagrams—vertical polarization.
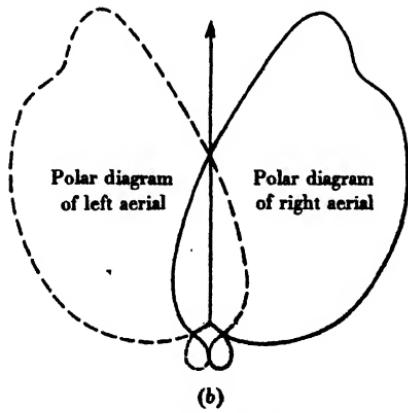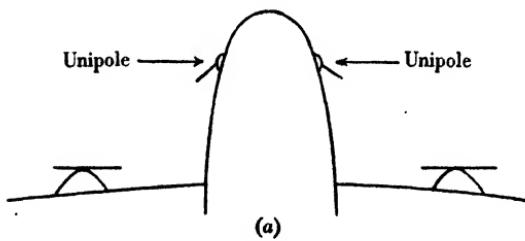
Fig. 85. Use of fuselage to obtain overlapping polar
diagrams—horizontal polarization.

the radiation in a forward direction so as to obtain a better polar diagram.

For the second method there are two types of polar diagram which have a form suitable for homing purposes:

(*a*) A diagram such as may be obtained from a driven element and reflector (fig. 30 (*b*)).

(*b*) The diagram obtained from a terminated $\frac{1}{2}\lambda$ rod (Chapter 8). The latter is shown in fig. 86.
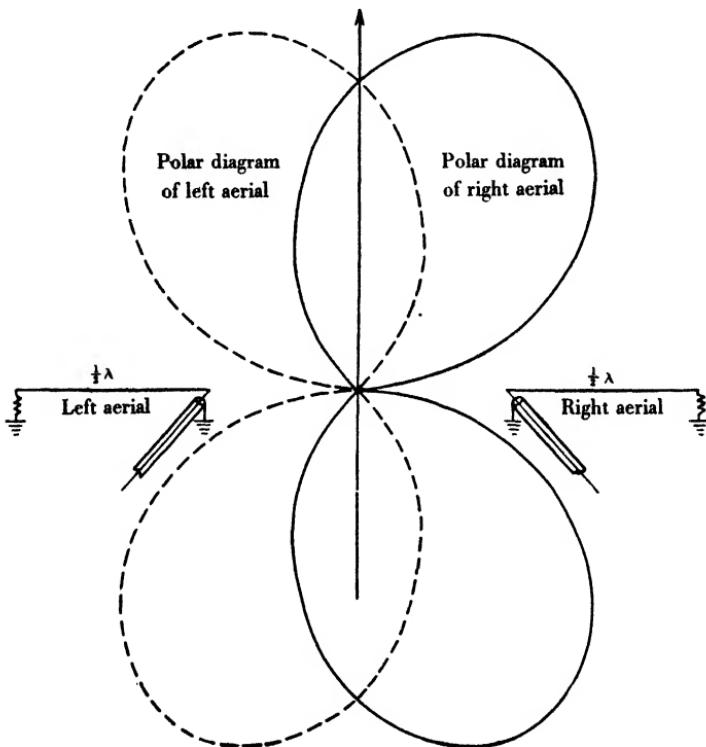
Fig. 86. Polar diagram of terminated $\frac{1}{2}\lambda$ rod used for homing.

In order to increase the band width over which homing is effective fat streamlined elements are used (e.g. $\frac{3}{4}$ in. major axis for 200 Mcyc./sec.). For some applications requiring a particularly large band width of operation three elements are used, reflector, driver and director, the reflector and director being adjusted to give optimum directivity at two different points in the band.

Generally speaking aerials which depend for their homing properties on their own free-space polar diagram are placed as far from the fuselage as possible, usually near the wing tips, or right in the nose of the fuselage.
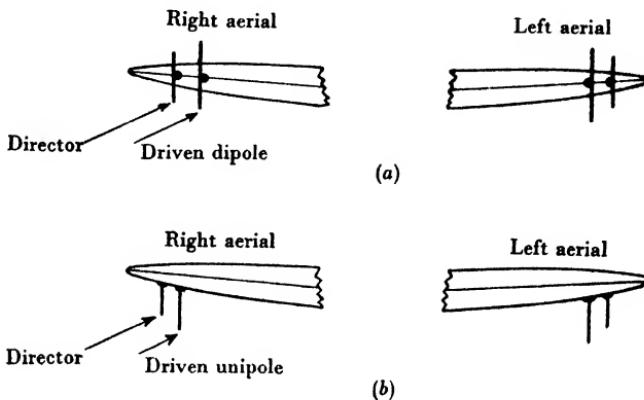
Fig. 87. Azimuth homing aerials mounted near wing tip.
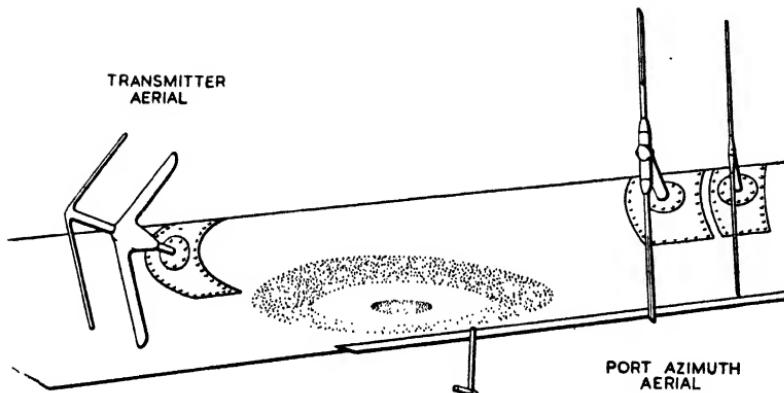
Fig. 88. A.I. aerials on Hurricane wing.

Two examples of aerials of this kind are shown in fig. 87. The actual form of the aerials will be seen from fig. 88, which shows some of the A.I. aerials on the wing of a Hurricane.

11·1·2. *Homing aerials with restricted azimuth coverage*

Yagi aerials are almost always used in order to provide narrow overlapping beams to give accurate homing for radar applications.

It will readily be seen that the accuracy is increased if the rate of change of signal with azimuth is great. For this purpose it is desirable to use a polar diagram in the form of a narrow main beam with smaller side lobes, and its use has already been described in Chapter 9 in connexion with the 'split' method of direction finding (see fig. 71 (*b*)). Two Yagi aerials are normally each inclined at an angle of 10–15° to right and left of the line of flight. These Yagi aerials are normally of all-metal construction as described in Chapter 10. A typical example mounted on a Liberator aircraft is shown in fig. 89 (Pl. II, p. 116). One such Yagi aerial is mounted under each wing for reception; another Yagi aerial with slightly wider coverage is mounted in the nose of the aircraft for transmission.

11·2. Aircraft aerials for producing narrow beams

In addition to the use of Yagi aerials for homing as described above, they have also been used, extended along the wings, to produce a narrow beam at right angles to the direction of flight for various purposes. Broadside arrays of dipoles using the fuselage as a reflecting screen have also been used. Such large aerial arrays produce considerable drag and reduce appreciably the speed of the aircraft.

11·3. 'All-round' aerials

It is much easier to produce a good 'all-round' polar diagram for vertical polarization than for horizontal polarization, since a vertical half-wave dipole produces a circular polar pattern in the horizontal plane, whereas a horizontal dipole produces a 'figure of eight' pattern. For vertical polarization a $\frac{1}{4}\lambda$ element mounted on the centre line of the fuselage, either above or below, about midway between the trailing edges of the wings and tail-plane leading edge, is usually satisfactory. For horizontal polarization a dipole bent in the form of a V with its two legs at right angles may be used. It is placed about $\frac{1}{2}\lambda$ from the fuselage. For larger aircraft it may be necessary to use two unipoles, one on each side of the fuselage, and bent at 45° to the line of flight as for homing aerials. In general, however, it is much more difficult to obtain a uniform diagram with horizontal polarization, and dips up to -10 db. must be expected.

11·3·1. *Propeller modulation*

For some applications, such as instrument landing by means of radio beams, it is essential to obtain steady signals. Now if the receiving aerial is near the propellers there is a tendency for the signals to be modulated as the propellers turn. Aerials for reception of instrument landing signals are therefore usually placed underneath the fuselage well back towards the tail.

11·3·2. *V.h.f. communication aerials*

By far the most important range of frequencies for aircraft communications in the metric band is the so-called v.h.f. band 100–150 Mcyc./sec., and it is only with this band with which we shall be concerned in this section.

For aircraft with metal skins unipoles have decided advantages over dipoles in that they may be fed at their base by means of a concentric line the outer conductor of which is earthed to the aircraft skin if this is metallic. A dipole fed at its centre would have to be mounted above the skin at a considerable distance, and since its free ends are high-voltage points they should not be near the skin. Since, however, a mounting of this type usually leads to difficult mechanical problems it is more convenient to employ unipoles.

For fairly low-speed aircraft the v.h.f. aerial usually consists of a hollow, streamlined, tapered steel mast of considerable cross-section (about 3 in. at base) in order to obtain an appreciable band width over which the aerial is effective. For aircraft not having a metal skin an earth must be provided. This may take the form of a sheet of metal or of dipole earth bars, the latter being usually preferred for structural reasons (see §7·1). These aerials may be used over a band of about 25 Mcyc./sec. fed directly by 45-ohm cable without introducing a mismatch of more than about $2\frac{1}{2}$ to 1.

For aircraft with speeds above 250 m.p.h. the drag of these aerials becomes serious. For this higher speed range whip aerials are used. These have much less drag, but their impedance varies much more rapidly with frequency and must be matched over the band by means of a series of matching units each covering about 10 Mcyc./sec. The whip aerials are made of tensile steel and are $\frac{1}{4}$ in. in diameter at the base and tapered to about $\frac{1}{16}$ in. They are about 27 in. long and are mounted so as to slope back from the vertical at an angle

which varies between about 5 to 20° according to the maximum level speed of the aircraft.

Even the whip aerials begin to have a serious drag at speeds over 450 m.p.h., and for these very high speeds special blade aerials have been developed at the Royal Aircraft Establishment. The blade has a major axis which tapers from about 2 in. at its base to ½ in. at its tip, its minor axis being five times less than its major axis. It is made from very high-grade duralumin, and so is both strong and light. The blade aerial may be matched over the band 100–150 Mcyc./sec. by the use of four matching units.

11·3·3. *Whip aerials for Gee, etc.*

For the reception of Gee signals in aircraft, whip aerials are commonly used(3). These whips are considerably longer than the v.h.f. whips, being cut to resonate as unipoles at about 45 Mcyc./sec. They consequently present much more severe mechanical problems.
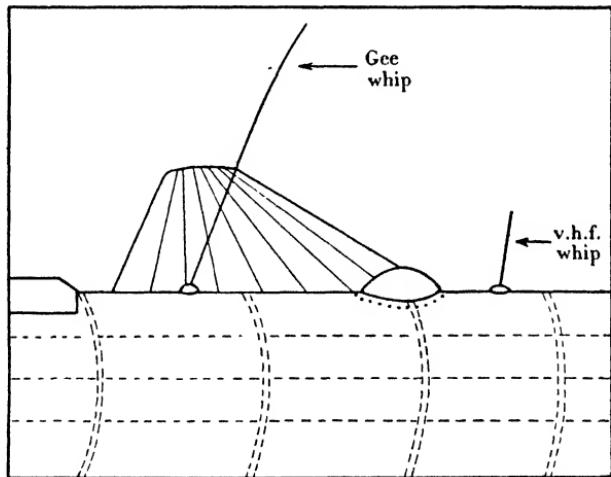
Fig. 90.   Gee whip aerial and v.h.f. communication aerial mounted on the fuselage of an aircraft.

These aerials were originally required to operate in the band 45–50 Mcyc./sec. and were matched to a 45-ohm cable by means of L-sections at five spot frequencies in the band. Later they were required to operate at five spot frequencies in the band 22–30 Mcyc./sec. and again were matched by 'preset' matching sections which were switched in for each spot frequency. Later, operation

was also required over a continuous band of frequencies from 50 to 85 Mcyc./sec. and a shorter whip was used, being matched over the band by a number of steps.

For G.H. operation(3) transmission at high-peak powers of the order of 20 kW. had to take place from a whip aerial, and this presented serious insulation difficulties. The frequencies used for these transmissions were in the band 20–30 Mcyc./sec., so resonant whips were somewhat long. Nevertheless, it was essential to use nearly resonant whips to make the base of the aerial a low-impedance point. The band was covered by a series of whips designed to screw into a plug fitted in the aircraft's skin. At a frequency of 22 Mcyc./sec. the whip aerial is becoming very clumsy, and its use at lower frequencies on aircraft is out of the question.

A typical installation showing a Gee whip and v.h.f. communication aerial mounted on the fuselage of an aircraft is shown in fig. 90.

11·4. 'Suppressed' aerials

When the speed of aircraft exceeds about 450 m.p.h. it is desirable to avoid all external aerials. For wooden aircraft, aerials are sometimes simply mounted inside the fuselage, but screening due to equipment carried by the aircraft is usually troublesome, and this is not a very good solution of the problem.

An interesting form of v.h.f. aerial has been developed at the Royal Aircraft Establishment in the form of a metallic foil covering part of the tail fin. It is only necessary for part of the fin to be non-metallic, and this may be made of wood or some plastic material. The arrangement is shown in fig. 91. Being in the form of a 'fat' unipole such an aerial would be expected to have a low rate of change of impedance with frequency, and it should be possible to match it easily over a wide band of frequencies. The foil aerial is fed by means of a concentric line, the outer being connected to the metallic part of the aircraft or to an earthing bar.

Another interesting form of 'suppressed' aerial has been used for homing aerials for the 'Rebecca' navigational system(1). The aerial consists simply of a driven element enclosed in a non-metallic wing tip. It is suitable for use with horizontal polarization on aircraft with metal wings. The metallic part of the wing acts as a reflector and produces the desired shape of polar diagram. The arrangement is shown in fig. 92.

The use of slot aerials cut in the metal skin of aircraft offers a neat solution of the problem of providing aerials which are flush with
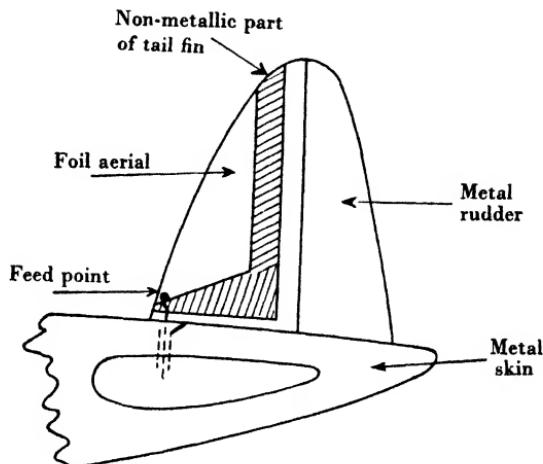
Fig. 91. Foil aerial mounted on tail fin.
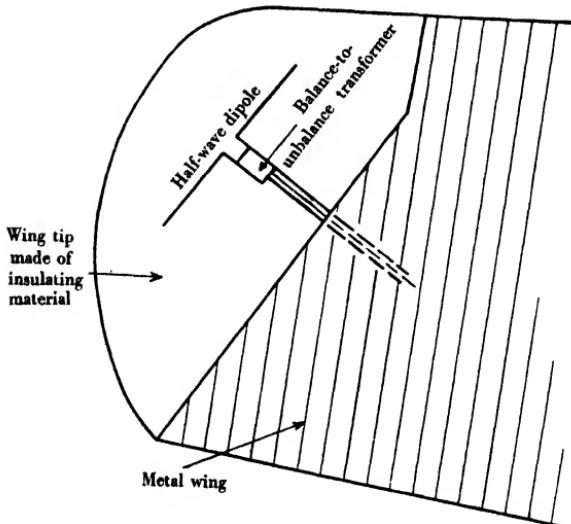
Fig. 92. Enclosed wing-tip homing aerial.

the fuselage. Some experiments on the use of this type of aerial have been successfully carried out, but further work is required before a wholly satisfactory arrangement is developed. The slots

are usually fed by cavities within the aircraft structure in order to screen them from internal metallic objects. It is clear that such aerials can only be used for quite short wave-lengths, otherwise the aircraft structure would be seriously weakened. Moreover, it will be almost essential to decide where such slot aerials are to be placed in the early stages of the design of aircraft in order that the structural design may take account of their presence.

REFERENCES

(1) Smith, R. A. *Radio Aids to Navigation*, Chapter 3. Cambridge University Press (1947).
(2) Russell, B. The design and positioning of aircraft radar aerials for metric wave-lengths. *J. Instn Elect. Engrs*, **93**, III A, 567 (1946).
(3) Ref. 1, Chapters 6 and 10.

Chapter 12

WIDE-BAND AERIALS

12·0. Introduction

For many purposes an aerial is required to operate only at a number of 'spot' frequencies. In this case it may be matched at each frequency. However, there are other applications which require operation over a wide band of frequencies and, moreover, require that the mismatch between the aerial and its feeders shall be small over the entire band. Clearly a compromise must be struck, as it is impossible to match an aerial over a wide band as accurately as may be achieved at a 'spot' frequency. In general the higher the standing-wave ratio which is tolerable the wider will be the band which can be covered.

We have already seen in Chapter 4 that the rate of variation of reactance of a dipole aerial with length, or with frequency, is greatest for thin aerials. For wide-band working, therefore, fat dipoles are used, and we shall first of all be concerned with their properties. We shall then consider various devices for extending the band width by compensating for the changes of reactance of the dipoles.

12·1. Equivalent circuits of half-wave and full-wave dipoles and definition of 'Q'

For a fairly thin half-wave dipole the variation of resistance near the resonant point is small compared with the variation of reactance, and we may represent the aerial by an equivalent series circuit as shown in fig. 93 (a). The value of R is just the input resistance R_0 of a thin half-wave dipole, 73·2 ohms. L and C are determined in terms of the length and radius of the aerial as shown below. The input impedance of the circuit shown in fig. 93 (a) at frequency f is given by

$$Z_i = R_0 + j \cdot 2\pi f L - j/2\pi f C. \tag{1}$$

The condition that Z_i should be purely resistive is

$$1/LC = 4\pi^2 f^2. \tag{2}$$

For a half-wave aerial $\quad Z_i = R + jX,$ $\tag{3}$

where in general for a fixed length $2l$ of the aerial both R and X are functions of f. We have seen, however, in Chapter 4 that when the aerial is thin and nearly resonant, i.e. for values of f near f_0, where

$$c/f_0 = 4l, \qquad (4)$$

then the variation of R with f is small compared with the variation of X. For values of f near f_0 the input impedance of the aerial is given approximately by

$$Z_i = R_0 + jX_0 - j\frac{\pi Z_0}{2}\frac{\Delta\lambda_1}{\lambda_1} \qquad (5)$$
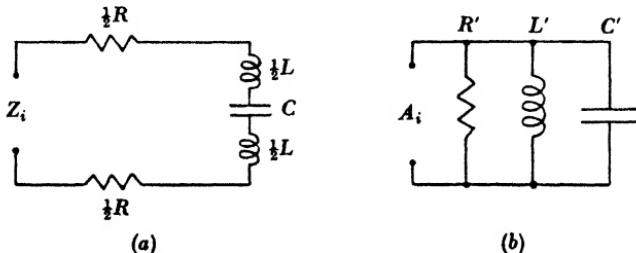
Fig. 93. Equivalent circuits of half-wave and full-wave dipoles.

[cf. equation (34)] of Chapter 4, where R_0 is the resistance and X_0 the reactance of an aerial exactly $\frac{1}{2}\lambda$ long, and $\Delta\lambda_1$ is the change in wave-length from the value λ_1 given by

$$\lambda_1 = 4l. \qquad (6)$$

Z_0 is the 'characteristic impedance' of the aerial and is equal to $120[\log_e 2l/a - 1]$.

Equation (5) may also be written in the form

$$Z_i = R_0 - j\frac{\pi Z_0}{2}\frac{\Delta\lambda}{\lambda_0}, \qquad (7)$$

where $\Delta\lambda$ is the change in wave-length from the value λ_0 which makes the dipole purely resistive. An alternative form of equation (7) is

$$Z_i = R_0 + j\frac{\pi Z_0}{2}\frac{\Delta f}{f_0}, \qquad (8)$$

where f_0 is the resonant frequency and $f = f_0 + \Delta f$. Equation (8) is only valid for values of Δf such that

$$\frac{\pi Z_0}{2}\frac{\Delta f}{f_0} \sim R_0. \qquad (6)$$

Now equation (1) may be written in the approximate form

$$Z_i = R_0 + j \cdot 4\pi L f_0 \frac{\Delta f}{f_0} \tag{10}$$

for small changes of frequency Δf, and we may compare (8) and (10) and thus find the value of L in the equivalent circuit. We have

$$L = Z_0/8f_0. \tag{11}$$

From (2) we have, therefore,

$$C = 2/\pi^2 Z_0 f_0. \tag{12}$$

Similarly, the input admittance of a full-wave dipole may be written as (cf. equation (42) of Chapter 4)

$$A_i = A_0 + j \frac{\pi Z_0}{Z_0'^2} \frac{\Delta f}{f_0}, \tag{13}$$

where f_0 is now the anti-resonant frequency, A_0 is the admittance when $f = f_0$, and $Z_0' = Z_0 - 120 \log_e 2$. A_0 is given by [Chapter 4, equation (44)]

$$Z_0'^2 A_0 = 196 \cdot 6 \text{ ohms.} \tag{14}$$

From the equivalent circuit in fig. 93 (b) we have

$$A_i = A_0 - \frac{j}{2\pi f L'} + j \cdot 2\pi f C', \tag{15}$$

where we have written $R' = 1/A_0$.

For small values of Δf

$$A_i = A_0 + 4\pi C' f_0 \frac{\Delta f}{f_0}. \tag{16}$$

Hence

$$C' = Z_0/4f_0 Z_0'^2 \tag{17}$$

and

$$L' = Z_0'^2/\pi^2 f_0 Z_0. \tag{18}$$

The band width over which a tuned circuit is effective is normally defined in terms of its 'Q'. For a series-resonant circuit the value generally taken for Q is $2\pi f L/R$. There are various physical interpretations of the quantity Q. For example, if an e.m.f. $l_0 e^{2\pi j f t}$ is applied to a circuit such as that shown in fig. 93 (a) the forced oscillation has current amplitude i given by

$$i = \frac{l_0}{\{R^2 + (2\pi f L - 1/2\pi f C)^2\}^{\frac{1}{2}}}.$$

The resonant value is l_0/R. If f_0 is the resonant frequency and f_1 the frequency for which i is reduced by 3 db., then

$$\frac{2|f_1 - f_0|}{f_0} = \frac{R}{2\pi f_0 L} = \frac{1}{Q}, \tag{19}$$

provided $Q \gg 1$. f_0/Q is therefore the band width over which the amplitude is not less than 3 db. below its maximum value.

For the half-wave aerial, applying these considerations, we have

$$Q_{\frac{1}{2}\lambda} = \frac{\pi Z_0}{4R_0} \tag{20}$$

$$= 1 \cdot 3 \left[\log_e \left(\frac{\lambda}{2a} \right) - 1 \right]. \tag{21}$$

For a full-wave aerial

$$Q_\lambda = 2\pi f_0 C'/A_0 = \frac{\pi Z_0}{2Z_0'^2 A_0} \tag{22}$$

$$= 0 \cdot 95 \left[\log_e \left(\frac{2\lambda}{a} \right) - 1 \right]. \tag{23}$$

Let us consider some numerical values. These are shown in table 14 for half-wave and full-wave aerials. From a circuit point of view then, even a fairly thin aerial is equivalent to a flatly tuned circuit.

Table 14. 'Q' of half-wave and full-wave aerials

λ/a	$Q_{\frac{1}{2}\lambda}$	Q_λ
50	2·9	2·7
100	3·8	3·4
500	5·8	5·0
1,000	6·6	5·6
5,000	8·8	7·2
10,000	9·6	7·8
50,000	11·7	9·5
100,000	12·7	10·1

The above considerations only strictly apply when Q is large, say greater than 10, since (10) is only valid for this condition. The values calculated from (21) and (23) are, however, good enough when $Q > 5$ to give a general impression of the behaviour of the aerial. For really wide-band aerials even smaller values of Q are required, and Q must then be defined in terms of the actual reactance and resistance variations themselves.

Suppose we have a half-wave aerial fed by a line of characteristic impedance R_0 which is matched to a generator of the same impedanee. At the resonant frequency the line will be matched at both its ends, but as the frequency is varied from the resonant value the power fed to the aerial for a given e.m.f. in the generator will fall. The power P dissipated in the aerial, if e is the e.m.f. of the generator, will be

$$P = \frac{e^2 R}{2\{(R_0 + R)^2 + X^2\}}, \qquad (24)$$

where $R + jX$ is the value of the impedance of the aerial at frequency f. P will be a maximum if $X = 0$ and $R = R_0$, i.e. when $f = f_0$. The maximum value will be given by

$$P_{\text{max.}} = e^2/8R_0. \qquad (25)$$

For frequencies greater and less than f_0, P will be reduced. There will be in general two frequencies, f_1', f_2', one on either side of f_0 for
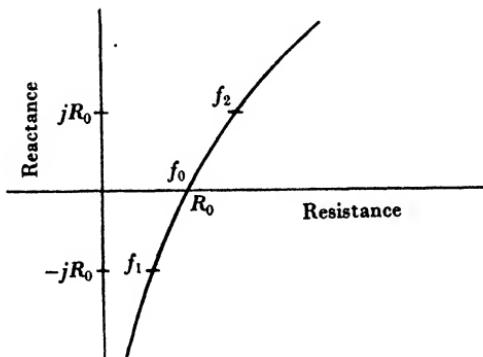
Fig. 94. Illustrating definition of 'Q' when both R and X vary.

which P is reduced by a factor of two and the Q of the aerial is sometimes defined as

$$f_0/|f_1' - f_2'|. \qquad (26)$$

This does not, however, line up with the usual definition for a tuned circuit, and a more usual definition of Q is as follows. Let f_1 and f_2 be the two frequencies on either side of f_0 for which $|X| = R_0$. Then Q is defined as

$$f_0/|f_1 - f_2|, \qquad (27)$$

which agrees with the usual circuit definition. This is illustrated in fig. 94. Another definition which is sometimes used is that Q is

the ratio of 2π times the mean stored energy to the energy radiated per cycle. These all lead to only slightly different values for Q.

12·2. Behaviour of fat cylindrical dipoles

The considerations of Chapter 4 have indicated that, so far, a satisfactory theoretical treatment of cylindrical dipoles has only been given for values of l/a such that $2\log_e 2l/a > 10$. It may be deduced, however, that the general behaviour of dipoles when l/a has smaller values is as follows. The resonant resistance of the half-wave dipole is not likely to fall much below 60 ohms. The anti-

———— Brown and Woodward   ◯ Essen and Olliver
× Rösseler, Vilbig and Vogt
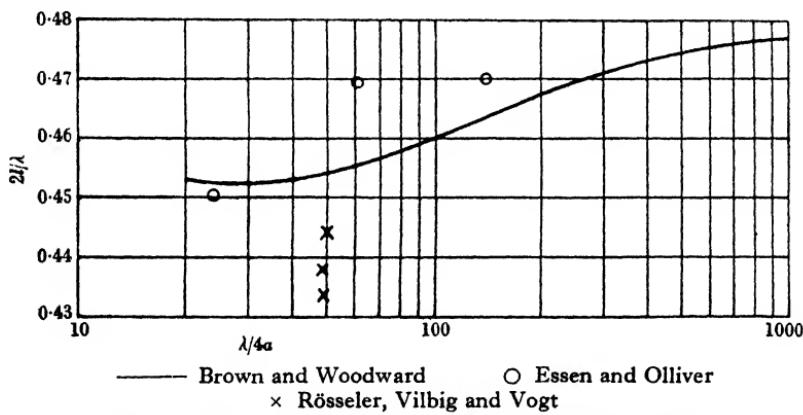
Fig. 95. Resonant length of fat, cylindrical, half-wave dipole.

resonant resistance will fall rapidly as l/a decreases and will tend towards the value of the resonant resistance. The values of l for which resonance and anti-resonance take place will decrease with l/a, and the maximum value of the reactance between the resonant and anti-resonant points will decrease rapidly. This behaviour is illustrated for the case of a fat spheroidal aerial for which a reasonably good theoretical treatment is available(4) (fig. 25).

For other dipoles we must have recourse to experiment. In fig. 95 we have summarized the scanty experimental information on the resonant length of fat, half-wave cylindrical dipoles. This information has been obtained from experiments by Brown and Woodward(1), Essen and Olliver(2), Rösseler, Vilbig and Vogt(3) already referred to. It is quite clear that when the impedance of a fat cylindrical dipole is measured by using a unipole above an earth

screen the impedance includes the effect of a large base capacity. This has the effect of increasing the resonant length over what it would otherwise be. This is well illustrated by the three points taken from the results of Rösseler, Vilbig and Vogt which were obtained with cylinders with plane and tapered ends, the most tapered cylinder giving the largest percentage decrease from the half-wave value. As we shall see later it is very desirable to use
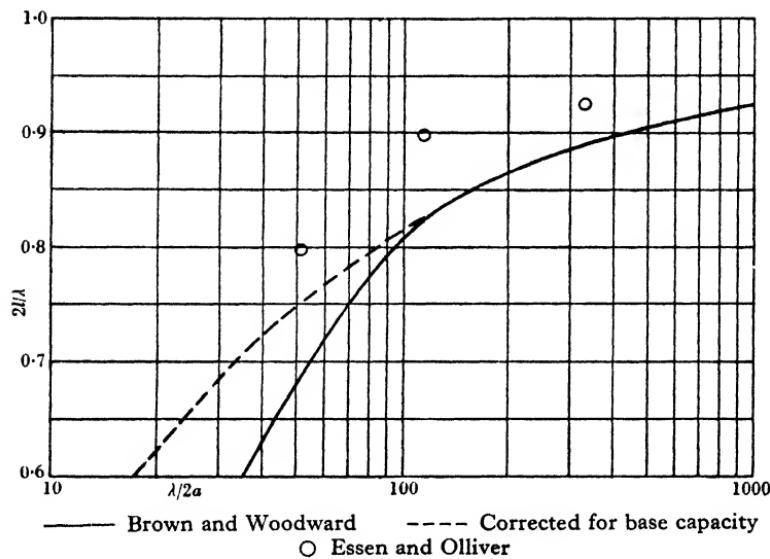
Fig. 96. Anti-resonant length of fat cylindrical dipole.

a tapered dipole to avoid the effect of base capacity which has an adverse effect on the band width of the dipole. Although general agreement between the different workers is obtained, it is clear from fig. 95 that for practical purposes it is desirable to measure the input impedance of the dipoles to be used to construct an aerial.

One interesting point which emerges from these measurements is that the resonant resistance of a fat dipole is almost constant, having a value between about 58 and 60 ohms for values of $\lambda/4a$ lying between 10 and 1000.

As the ratio $\lambda/4a$ decreases, the input resistance of a dipole at its anti-resonant point decreases rapidly to a value of a few hundred ohms. Such a dipole may be fed directly with an open-wire transmission with very small mismatch, and this has wide applications in building of arrays for wide-band working. From this point of

view the fat full-wave dipole is more important than the half-wave dipole. Moreover, the shortening of the dipole is considerable and this has mechanical advantages. Values for the anti-resonant length and input resistance of fat dipoles are shown in figs. 96 and 97 as obtained by Brown and Woodward[1] and Essen and Olliver[2]. In fig. 96 is shown, by means of the dotted curve, the effect of correcting for the base capacity, which was obtained by electrostatic considerations. When this is taken into account a considerable
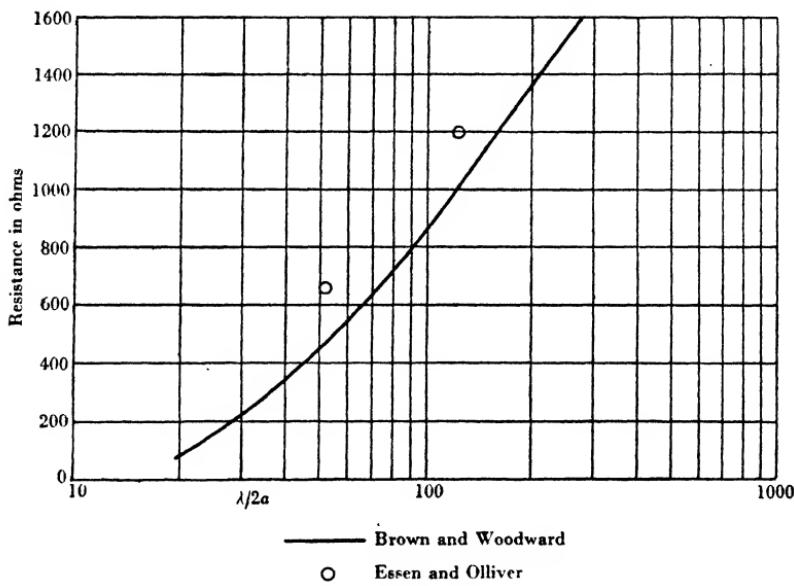
Fig. 97. Anti-resonant resistance of fat cylindrical dipole.

increase takes place in the anti-resonant length of a fat dipole. The maximum value of the reactance attained between the resonant and anti-resonant points is shown in fig. 98. This quantity is of interest as it determines to a large extent the mismatch which will take place if the frequency is varied over a wide band. The values shown are taken from experimental curves by Brown and Woodward[1] and Essen and Olliver[2]. They are only intended to illustrate the general behaviour of fat full-wave dipoles. It is quite clear that, apart from any fundamental effect of the gap at the centre, the base capacity will vary as the gap is varied, and for fat dipoles this is bound to play an important part in determining the reactance. What is wanted is a whole series of measurements with varying gaps in the middle

in order to separate out the effect of the base capacity from more fundamental effects. So far as is known, such measurements do not yet exist.

In fig. 99 are shown the variation with length of the input resistance and reactance of two very fat cylindrical dipoles, one measured by Brown and Woodward[1] and the other by Essen and Olliver[2]. In the latter measurements the length of the dipole was fixed and the frequency varied so that the ratio $\lambda/2a$ varied. Moreover, the inner ends of the dipole were coned in order to reduce
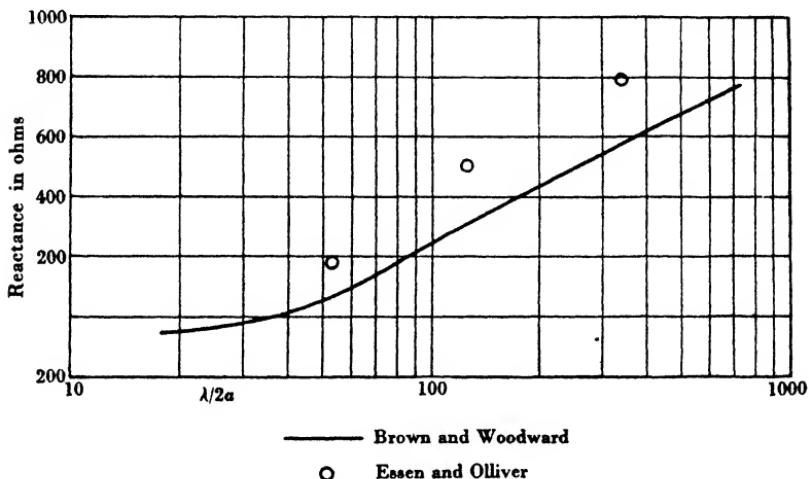
Fig. 98. Maximum reactance between resonant and anti-resonant points.

base capacity, and this largely accounts for the difference in behaviour. The reactance of the dipole measured by Brown and Woodward is wholly negative due to the base capacity. This also has the effect of reducing the input resistance.

12·3. Cage dipoles

Fat cylindrical dipoles are usually too clumsy to use except at very short wave-lengths, and for the longer wave-lengths in the metric band are frequently replaced by cage dipoles. These cages are made up from wires arranged in a form having a circular or square section. They are normally coned at the inner end to reduce base capacity. Measurements on unipoles of this form have been made by Rösseler, Vilbig and Vogt[3], and by Lutkin, Cary and

Harding(5). The characteristic impedance of such dipoles may be obtained by measuring the rate of change of reactance near the resonant or anti-resonant points and interpreting the results in
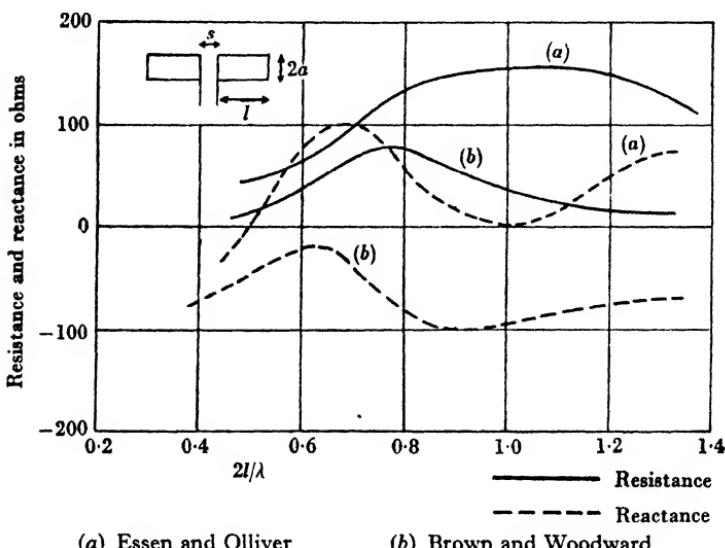
(a) Essen and Olliver   (b) Brown and Woodward

Fig. 99. Variation of input resistance and reactance
with length of two fat cylindrical dipoles.
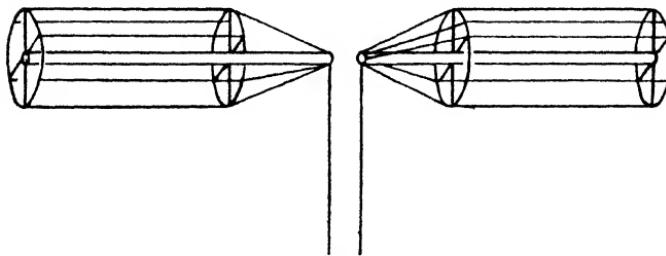
Fig. 100. Cage dipoles.

terms of equations (8) or (13). The values obtained for cage dipoles are found to be of the same order as those for solid dipoles having the same form and cross-section as the cage. The behaviour of the dipoles is then largely determined by their value of Z_0 and by the base capacity at the input. We shall later discuss the building of arrays from such dipoles. An example of the form of rigid construction used by Lutkin, Cary and Harding is shown in fig. 100.

The dipole is supported by a stout central conductor which gives it its mechanical rigidity but contributes little to its electrical properties.

12·4. Conical aerials

We have seen that coning the ends of a cylindrical dipole has the effect of reducing the base capacity. The use of dipoles of conical construction throughout their length has, however, some advantages and has been used by a number of workers, including Lutkin, Cary
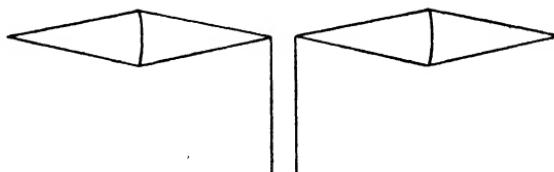
Fig. 101. Double conical dipole.

and Harding, who used cages in conical form as well as cylindrical cages with tapered ends. The theory of conical dipoles has been given by Schelkunoff[6] and has been discussed in Chapter 4. The analysis, however, only applies to cones having a fairly narrow angle at the vertex. The characteristic impedance of such cones is given by (see Chapter 4)

$$Z_0 = 120 \log_e (\cot \tfrac{1}{2}\alpha), \tag{28}$$

where α is the 'semi-vertical angle' of the cone.

In common with all 'fat' dipoles the rate of variation with frequency of reactance and resistance of conical dipoles with fairly large values of α is small compared with that of thin cylindrical dipoles (see, for example, Ref. (2)). It is this property and their convenient shape for feeding without introduction of appreciable base capacity that has made them popular with designers of wide-band aerials. Shortening due to the effect of the base is also an advantage.

The theoretical considerations have indicated that it is desirable to have an aerial 'fat' where the current is high. For full-wave dipoles the double cone as shown in fig. 101 has been used as it has just this property. Schelkunoff[6] has shown that the characteristic impedance of such a dipole is given by

$$Z_0 = 120 \log_e 2l/a, \tag{29}$$

where a is the maximum radius.

12·5. Conical feed for unipoles

We have seen that the conical aerial largely eliminates the base capacity associated with a cylindrical aerial. It also achieves another important purpose, namely, that of eliminating discontinuities at the feed point. A cylindrical aerial fed as in fig. 52 (a) introduces a discontinuity at the connexion of the inner conductor to the unipole unless its radius is the same as that of the unipole. Such discontinuities seem to introduce a rapid change of reactance with frequency which is just what is to be avoided in wide-band aerials. If, however, we use a tapered feed the discontinuity is avoided in
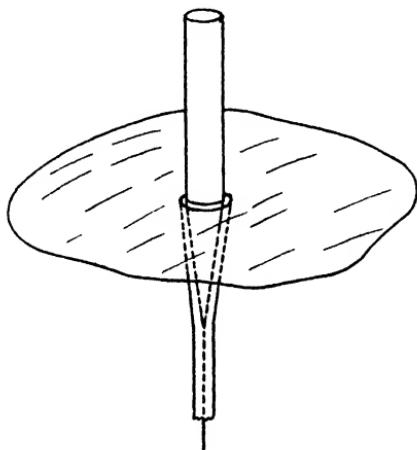
Fig. 102. Tapered feed for cylindrical unipole.

the case of a cylindrical unipole. This method of feeding is illustrated in fig. 102. The concentric line has the diameter of its inner conductor increased till it is equal to that of the unipole. The diameter of the outer conductor is also increased to maintain the value of the characteristic impedance of the line constant. The length of the tapered section should be about $\frac{1}{2}\lambda$ at the lowest frequency to be used.

With this arrangement cylindrical dipoles have almost as good band-width properties as any other form. A slight improvement may be obtained by rounding off the top into a spherical cap when the length is only three or four times the diameter.

The behaviour of two cylindrical unipoles with flat tops fed in this way is shown in fig. 103. The components of the admittance as

well as of the impedance are shown, as a knowledge of the behaviour
of these is frequently more desirable from a matching point of view
than the components of the impedance. Note that the conductance
has a maximum near the first resonant point. The value of this
maximum does not change much with diameter, but for thin
cylinders it falls more rapidly on either side than for fat cylinders.
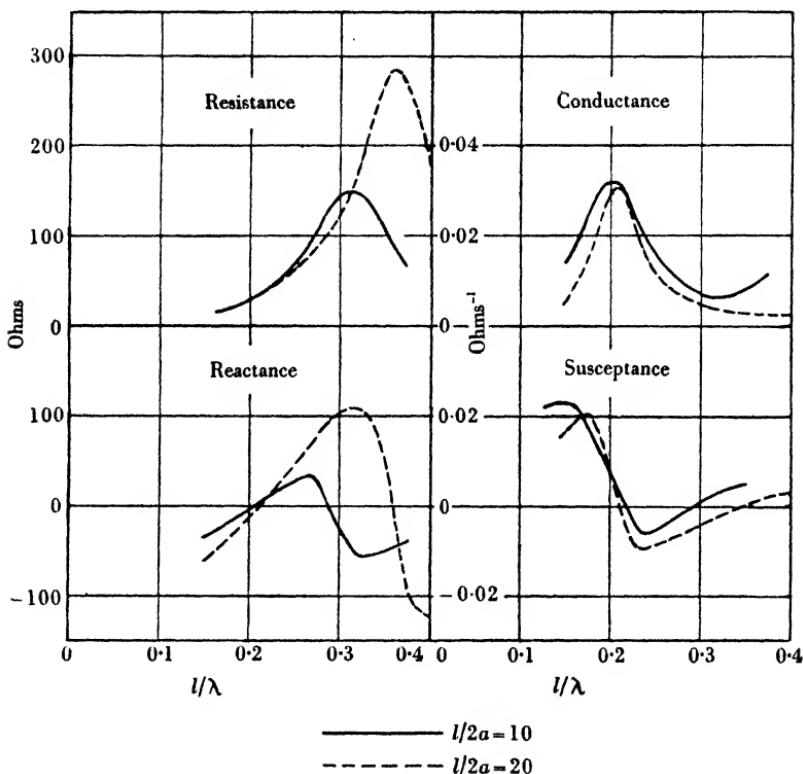
Fig. 103. Impedance and admittance of cylindrical unipoles with tapered feed.

12·6. Some special forms of wide-band dipoles and unipoles

A great deal of ingenuity has gone into the design of special forms
of dipoles and unipoles having wide-band properties. Most of this,
however, boils down to having a fat element and no discontinuities
or lumped impedances in the feeding arrangements. The well-
known R.C.A. television antenna, consisting of spheroidal elements
fed from concentric lines having special ends, is a good example(7).

12·7. Improvement of band width by stub compensation

The general form of variation of admittance of a dipole is similar to that shown in fig. 103. We note that the conductance has a maximum (stationary) value near the half-wave resonance, and that the susceptance changes in an approximately linear fashion from positive to negative values. Now the susceptance of a $\frac{1}{4}\lambda$ closed stub does precisely the opposite. For a frequency change Δf from the value f_0 for which the stub is $\frac{1}{4}\lambda$ long the susceptance is given by

$$S = -jz_0^{-1}\cot\left[\frac{\pi}{2}\left(1+\frac{\Delta f}{f_0}\right)\right], \tag{30}$$

where z_0 is the characteristic impedance of the stub line. If Δf is fairly small we have

$$S = \frac{\pi}{2z_0}\frac{\Delta f}{f_0}. \tag{31}$$

Near the first resonance the admittance A of a dipole may be written

$$A_d = C - jk\frac{\Delta f}{f_0}, \tag{32}$$

where f_0 is the frequency which makes the aerial resonate. If we place a closed stub $\frac{1}{4}\lambda$ long at $f=f_0$ across the dipole we have the total admittance given by

$$A = C + j\left(\frac{\pi}{2z_0} - k\right)\frac{\Delta f}{f_0}. \tag{33}$$

By choosing the characteristic impedance of the stub so that

$$\pi/z_0 = 2k, \tag{34}$$

we remove the first-order variation in susceptance. Moreover, since the conductance C has a stationary value near $f=f_0$ the variation of admittance will be small over a band width determined by the extent of the flat top of the maximum of the conductance curve and the linear part of the susceptance curve. This effect will be greatest for fat dipoles. If a line of characteristic impedance approximately equal to the reciprocal of the maximum conductance is available a good match will be achieved. If the stub is $\frac{1}{4}(2n+1)\lambda$ long at $f=f_0$ we must have

$$\pi/z_0 = 2k/2n + 1. \tag{35}$$

If the slope of the susceptance curve is positive, as for a full-wave dipole, a more elaborate procedure must be adopted. If we measure the admittance A' at a distance d along the line from the dipole it will be of the form

$$A' = C_0 + C_1 \Delta f + C_2 (\Delta f)^2 + \dots + j S_0 + j S_1 (\Delta f) + j S_2 (\Delta f)^2 + \dots \quad (36)$$

Now there will be certain values of d for which $C_1 = 0$ (usually about $\frac{1}{4}\lambda$ apart). At some of these points S_1 will be positive and at others negative. If we now choose an appropriate value of d and
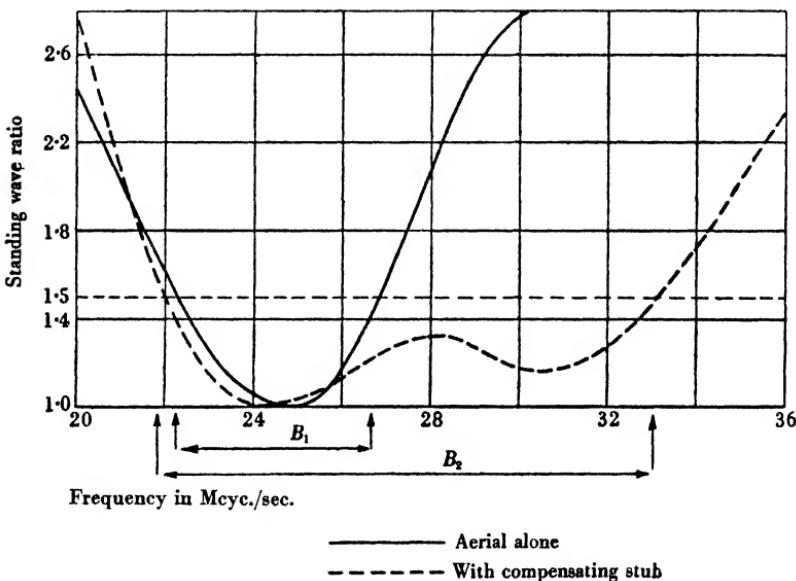
Fig. 104. Example of increase of band width by use of a compensating stub.

place a stub across the line at this point so as to compensate the variation $j S_1 \Delta f$ as explained above, the first-order terms may again be removed. An example of how the band width may be increased by this procedure is shown in fig. 104. It has been shown by Westcott and Goward[8] that the second-order terms may be reduced by a further arrangement of compensating stubs. Such compensated elements have been used for feeding paraboloidal reflectors over a wide band of frequencies centred on 600 Mcyc./sec. (see Chapter 14). When the slope of the susceptance curve is positive as l is increased, compensation by means of a parallel stub is impossible.

A series stub will, however, now effect the compensation, but is not so convenient as the parallel stub.

12·8. Wide-band matching transformers

When the above devices have been applied in order to make the impedance of the aerial itself as constant as possible over the required band of frequencies, there still remains the problem of matching the aerial to a line, unless the midpoint conductance happens to match. Most matching transformers operate strictly
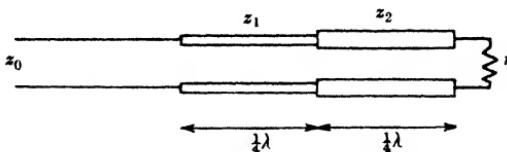
Fig. 105. Double $\frac{1}{4}\lambda$ transformer.

at only a single frequency. For example, the quarter-wave transformer turns a resistive load into another resistive load only when the frequency is such that it is exactly $\frac{1}{4}\lambda$ long. For a small change in frequency Δf from this value f_0, if r is the original load, we have for the impedance Z at the end of the transformer at frequency $f_0 + \Delta f$,

$$Z = \frac{Z_0^2}{r}\left[1 + j(r - Z_0)\frac{\pi}{2}\frac{\Delta f}{f}\right]. \qquad (37)$$

The first-order change of Z with frequency is only zero provided $Z_0 = r$.

12·8·1. Double $\frac{1}{4}\lambda$ transformer

If, however, we use two quarter-wave transformers in series the first-order variation of impedance with frequency may be made zero.

Let z_1 and z_2 be the characteristic impedances of the two transformers which are each $\frac{1}{4}\lambda$ long when $f = f_0$ (fig. 105). Suppose we wish to match a resistive load r to a line of characteristic impedance z_0 over a band centred at $f = f_0$. When $f = f_0$ the input impedance will be given by

$$Z = r z_1^2 / z_2^2. \qquad (38)$$

We must therefore choose the ratio of z_1/z_2 so that

$$z_0 z_2^2 = r z_1^2. \qquad (39)$$

To determine the input impedance when $f = f_0 + \Delta f$ let us write the electrical length of each section equal to $\frac{1}{2}\pi + \epsilon$. We now make a straightforward application of the transmission-line formula

$$Z = Z_0 \frac{z \cos\theta + jZ_0 \sin\theta}{Z_0 \cos\theta + jz \sin\theta} \tag{40}$$

for the input impedance of a line of characteristic impedance Z_0 and electrical length θ terminated by a load of impedance z, to calculate first the impedance at B and then at A, approximating to the first order in ϵ. After some rather tedious algebra the result is

$$Z = \frac{z_1^2 r}{z_2^2} \left[1 - j\epsilon \left(\frac{z_2}{z_1 r} - \frac{r}{z_2^2} \right)(z_1 + z_2) \right]. \tag{41}$$

The first-order term in ϵ and therefore in Δf may be made zero by choosing z_1 and z_2 so that

$$\frac{z_2}{z_1 r} = \frac{r}{z_2^2}. \tag{42}$$

Combining (42) with (39) this may be written as

$$\frac{z_0}{z_1} = \frac{z_2}{r}. \tag{43}$$

If we write k for the transformation ratio, i.e. $k = r/z_0$ we have

$$z_1 = k^{\frac{1}{4}} z_0, \quad z_2 = k^{\frac{3}{4}} z_0. \tag{44}$$

This analysis may be extended to a larger number of sections. It is best to have an even number with

$$\frac{z_0}{z_1} = \frac{z_2}{z_3} = \frac{z_3}{z_4} = \dots = \frac{z_{2n}}{r}. \tag{45}$$

For most practical purposes two is sufficient.

12·8·2. *Tapered transformers*

When the number of $\frac{1}{4}\lambda$ sections becomes large a transformer for which (45) is satisfied corresponds approximately to a line whose characteristic impedance varies exponentially from Z_0 to r. This has led to the idea of the exponential line transformer, which has been discussed theoretically by Wheeler[9]. It is not, in fact, essential to use a line whose characteristic impedance varies exponentially, and a line with uniform taper one or two wave-lengths long makes quite a good aperiodic transformer[10].

A theoretical discussion of the best form of taper has been given recently by Gent and Wallis, who also give a good list of references to the literature on the subject(11).

A practical form of tapered transformer with good aperiodic properties for passing from a 300-ohm four-wire transmission line to a two-wire 600-ohm line has been described by Lutkin, Cary and Harding(5) and is illustrated in fig. 106.
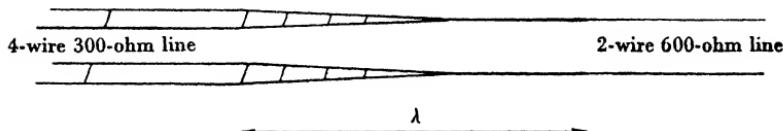
Fig. 106. Tapered transformer.

12·9. Arrays of wide-band elements

Arrays of wide-band elements have, of course, been built for television, and the problems of feeding them and maintaining a good match over a long length of transmission line for a fairly wide band of frequencies have been described by Cork and Pawsey(12).

The problem of designing aerials to cover a very wide band of frequencies and avoiding serious reflexions arose in connexion with the Gee navigational system. These aerials were required to cover three bands of frequencies: 20–30, 40–50 and 50–85 Mcyc./sec. A separate aerial system was used for each band. These aerials, and the problems of feeding them, have been discussed by Lutkin, Cary and Harding(5). Branch feeding was used in all cases in order to maintain the relative phase of the currents in the elements of the arrays when the frequency was varied. This form of feeding is illustrated in fig. 51. The elements themselves were matched to the line feeding them by choosing the dimensions of the elements to fit a 600-ohm line. When two branches were joined the junction was fed by a 300-ohm four-wire line. When two of these were joined one should have used a 150-ohm line. An attempt was made to use such a line made up of six wires but proved unsuccessful owing to discontinuities. An aperiodic tapered transformer as shown in fig. 106 was therefore used to transform up to a two-wire 600-ohm line before the junction was made and therefore 300-ohm four-wire line was used again.

The transmitting aerials for Gee are frequently remote from the transmitter building—often half a mile away. The problem of matching out standing waves produced by insulators, gantry supports, etc., at a spot frequency is relatively simple but over a band is very difficult. Any regular recurrence of a length along the line must be avoided, otherwise there may be a frequency in the band at which cumulative build-up of small loads in parallel will take place and may produce a very high standing-wave ratio on the line. Individual supports must be compensated in a wide-band fashion. Various methods of doing this have been described by Lutkin, Cary and Harding(5). These mostly depend on changing the characteristic impedance of the line over a small section of its length to produce a localized reactance which compensates that due to the insulator or support.

REFERENCES

(1) Brown, G. H. and Woodward, O. M. Experimentally determined characteristics of cylindrical antennas. *Proc. Inst. Radio Engrs*, *N.Y.*, **33**, 257 (1945).

(2) Essen, L. and Olliver, M. H. Aerial impedance measurements. *Wireless Engr*, **22**, 589 (1945).

(3) Rösseler, G., Vilbig, F. and Vogt, K. Über das elektrische Verhalten von Vertikalantennen in Abhängigkeit von ihrem Durchmesser. *T.F.T.* **28**, 170 (1939).

(4) Chu, L. J. and Stratton, J. A. Forced oscillations of a prolate spheroid. *J. Appl. Phys.* **12**, 241 (1941).

(5) Lutkin, F. E., Cary, R. H. J. and Harding, G. N. Wide-band aerials and transmission lines for 20–85 Mc./s. *J. Instn Elect. Engrs*, **93**, IIIA, 552 (1946).

(6) Schelkunoff, S. A. Theory of antennas of arbitrary size and shape. *Proc. Inst. Radio Engrs*, *N.Y.*, **30**, 493 (1942).

(7) Lindenblad, N. E. *Television Transmitting Antenna for Empire State Building*, p. 19. Radio at UHF. R.C.A. Institute Technical Press (1940).

(8) Westcott, C. H. and Goward, F. K. Design of wide-band aerial elements for 500–600 Mc./s. ground radar. *J. Instn Elect. Engrs* (in press).

(9) Wheeler, H. A. Transmission lines with exponential taper. *Proc. Inst. Radio Engrs*, *N.Y.*, **27**, 65 (1939).

(10) Starr, A. T. The non-uniform transmission line. *Proc. Inst. Radio Engrs*, *N.Y.*, **20**, 1052 (1932).

(11) Gent, A. W. and Wallis, P. J. Impedance matching by tapered transmission lines. *J. Instn Elect. Engrs*, **93**, IIIA, 559 (1946).

(12) Cork, E. C. and Pawsey, J. L. Long feeders for transmitting wide side bands with reference to the Alexandra Palace aerial feeder system. *J. Instn Elect. Engrs*, **84**, 448 (1939).

Chapter 13

SLOT AERIALS

13·0. Introduction

The most fundamental advance in aerial practice during recent years is undoubtedly the introduction of slot aerials. It seems as though these were used both in Britain and in the U.S.A. before their true significance was fully understood. It was H. G. Booker[1] who first appreciated the true relationship between slots and dipole aerials and the connexion of this with Babinet's principle as used in optics. In order to understand the theory of slot aerials we shall first have to discuss Babinet's principle and Booker's important extension of it.

13·1. Babinet's principle

Suppose we have a plane, black (perfectly absorbing), thin screen having in it a number of holes of any shape and size, and suppose light from a source is incident on the screen. This will produce a light pattern behind the screen. Now suppose we replace the screen by a complementary screen in which the holes in the original are replaced by black material and the rest is open. A new light pattern will now be produced. Babinet's principle states that if the two light patterns behind the screens are superimposed the resulting pattern will be just that which would be obtained if no screen were present. When the sizes of the holes are large compared with the wave-length and geometrical optics may be applied, the principle is obvious, as shadows and bright regions in the two patterns just fit together. The principle is, however, also true when diffraction is taken into account, and still holds when the dimensions of the holes are comparable with the wave-length.

In this statement no mention has been made of polarization. This does not arise if unpolarized light is used. In application of the principle to radio waves, however, polarization must be considered and a more exact statement of Babinet's principle is necessary. This has been given by H. G. Booker[1]. It is more usual in radio applications to deal with reflecting screens rather than absorbing screens, and Booker stated the principle in terms of the former.

The main new point that was introduced was that in passing to the complementary screen it is necessary to interchange the electric and magnetic vectors at each point. This implies a change of source to what may be called a conjugate source. For example, a small Hertzian dipole and a small current loop or magnetic dipole (fig. 12) are conjugate sources. For a train of plane-polarized plane waves the conjugate is a train of plane waves with the polarization turned through 90°.

Booker's extension of Babinet's principle may be stated as follows. Let radiation from a source or group of sources fall on a screen S_1 and produce an electric field $\mathbf{E}_1$ at a point P behind the screen. Now let the sources be replaced by their conjugate sources and let screen S_1 be replaced by S_2, its complementary screen. Let $\mathbf{E}_2$ be the electric field produced at P. Then if $\mathbf{E}_{01}$ and $\mathbf{E}_{02}$ are respectively the electric fields produced at P by the sources and their conjugates, in the absence of the screen,

$$\frac{\mathbf{E}_1}{\mathbf{E}_{01}} + \frac{\mathbf{E}_2}{\mathbf{E}_{02}} = 1. \tag{1}$$

For plane waves incident on the screen, instead of interchanging the electric and magnetic vectors, we may turn the complementary screen through a right angle.

A similar statement may be made regarding the reflected radiation. Let $\mathbf{E}_1'$ and $\mathbf{E}_2'$ be the reflected fields corresponding to $\mathbf{E}_1$ and $\mathbf{E}_2$, at a point P' in front of the screen and $\mathbf{E}_{01}'$ and $\mathbf{E}_{02}'$ the fields reflected from a perfect screen; then we have

$$\frac{\mathbf{E}_1'}{\mathbf{E}_{01}'} + \frac{\mathbf{E}_2'}{\mathbf{E}_{02}'} = 1. \tag{2}$$

The proof of this form of Babinet's principle involves rather fundamental electromagnetic theory and will not be given here (see Refs. (2, 9)).

As an example of Babinet's principle consider a screen in which an array of slots has been cut as shown in fig. 107 (*a*). The complementary screen (fig. 107 (*b*)) will consist of an array of strips. These will act as parasitic dipoles, and such an array will act as a reflecting screen when the wave-length is nearly equal to half the length of each strip (the reflexion will only be perfect for an infinite array) otherwise the screen will be transparent. By Babinet's principle

the complementary screen (a) will act as a reflector when the wave-length is not nearly equal to half the length of the slot, but when this is so it will become transparent. Note that the screen (b) has been turned through a right angle relative to (a). Clearly the electric field must be *along* the strips but *across* the slots.

(a)   (b)
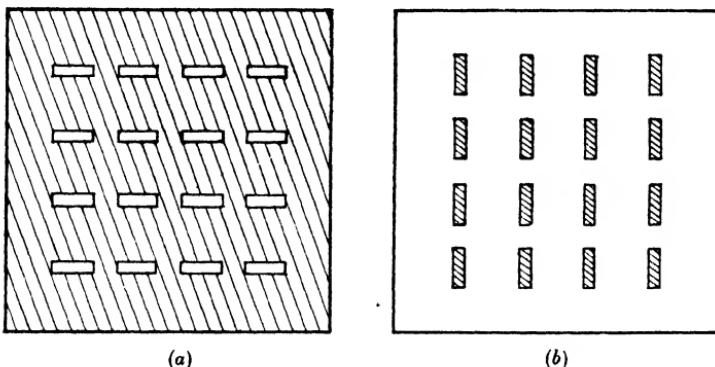
Fig. 107. Complementary screens.

13·2. The resonant slot

If we have only a single half-wave slot the complementary screen will consist of a half-wave strip. This will behave just like a half-wave cylindrical dipole of a certain radius. (The value is one-quarter
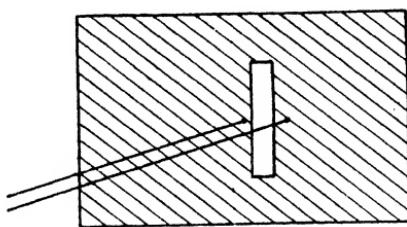
Fig. 108. Slot fed by parallel wire transmission line.

the width of the strip; see equation (14).) We should therefore expect a close relationship between the properties of a half-wave slot in an infinite sheet of metal and a half-wave dipole. This is, in fact, so. The slot radiates like an aerial and has a polar diagram similar to that of a half-wave dipole. The polarization is, however, turned through 90°, the electric field being at right angles to the length of

the slot. The slot may be fed by means of a generator at its centre producing a potential difference between its edges. In práctice this frequently takes the form of a parallel wire transmission line as shown in fig. 108. Moreover, there is a definite relationship between the input impedance of the strip dipole and that of the slot. We shall follow Booker's method of deriving this relationship.

Clearly in any electromagnetic field we may interchange everywhere electric and magnetic quantities. We shall, of course, have to replace perfect conductors of electricity by 'perfect conductors of magnetism'. Such materials do not exist in nature but are useful for purposes of argument. They have the property that the magnetic field at the surface must be normal.
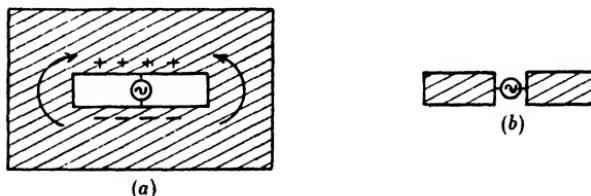
Fig. 109. Strip dipole and complementary slot.

Let us consider then the electromagnetic field due to a strip dipole (fig. 109 (b)). This will be just that for a half-wave dipole (Chapter 3, equation (10)), and the input impedance will be that given by equation (34) of Chapter 4 with (cf. equation 14)

$$Z_0 = 120[\log_e(8l/\omega) - 1].\tag{3}$$

Now let us interchange all electric and magnetic quantities. The dipole will become one made of 'perfectly conducting magnetic material' fed by a magnetic generator which we may regard as a current loop having two generators each driving current round the same way. This is illustrated in fig. 110 (a). The field produced will be the same as that of an electric dipole with E and H interchanged. The magnetic field at the surface of the strip will be normal and in opposite directions on the two sides. At other points in the plane of the strip the electric field will be normal and, of course, continuous. This may be seen from interchanging E and H for the original dipole. Now place the strip of perfect magnetic conductor in the slot in the complementary screen (fig. 110 (b)), which it will fit perfectly. The

field will be unchanged, since E was everywhere normal to the complementary screen. Now reverse the generator on the right of the screen. This may be done as the screen is now complete and the fields to right and left are independent. We have now the conditions shown in fig. 110(c). H is now continuous across the 'perfect magnetic material', and this may now be removed leaving the condition as shown in fig. 110(d) in which the generators are effectively in parallel. This corresponds to a slot fed across its width at the centre. We have not changed the fields except for a reversal on one side of the screen, and the original interchange of E and H.
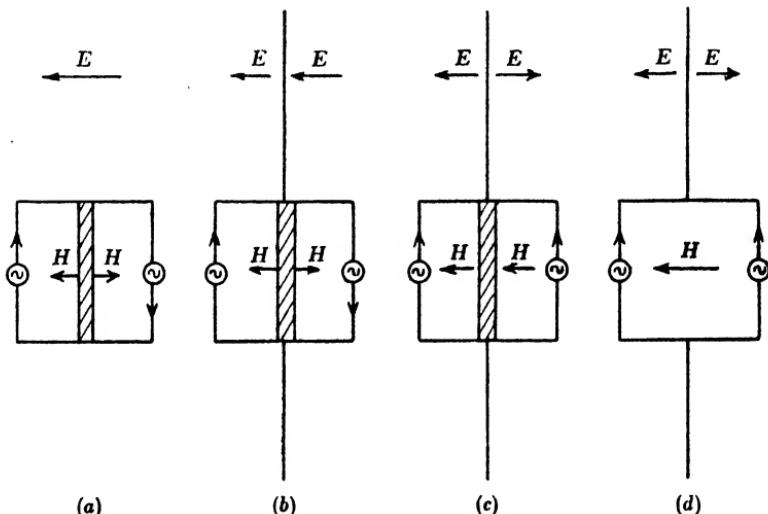
Fig. 110. Passage from dipole to slot.

The slot aerial therefore, with this reversal taken into account, produces exactly the same field pattern as the strip dipole but with magnetic and electric vectors interchanged.

We may now proceed to calculate the input impedance of the slot. Let E_1 and H_1 be the electric and magnetic fields associated with the strip dipole. Then in the m.k.s. units we have

$$E_1/H_1 = \zeta_0, \tag{4}$$

where ζ_0 is called the impedance of free space and is equal to 120π or 377 ohms(3). Let E_2 and H_2 be the electric and magnetic fields produced by the slot. On the left of the screen if we have $E_1 = H_2$ then $E_2 = \zeta_0^2 H_1$, since

$$E_2/H_2 = \zeta_0. \tag{5}$$

Now let e_1 and i_1 be the voltage and current at the input to the dipole and e_2 and i_2 the corresponding quantities at the input to the slot. Then

$$e_1 = \int_{c_1} E_1 ds, \tag{6}$$

$$i_1 = 2 \int_{c_1} H_1 ds, \tag{7}$$

$$e_2 = \int_{c_2} E_2 ds, \tag{8}$$

$$i_2 = 2 \int_{c_1} H_2 ds, \tag{9}$$
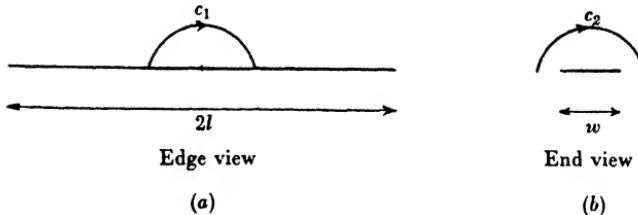
Fig. 111. Contours c_1 and c_2.

where c_1 is a small semicircle in a plane at right angles to the plane of the dipole or slot and parallel to the length of either, its centre being at the midpoint. c_2 is a similar contour but its plane is at right angles to the length of the strip. This is illustrated in fig. 111. (6) and (8) follow simply from the definition of e_1 and e_2. (7) follows from the relation between magnetic field and flux of current through a closed contour, using symmetry considerations. (9) also follows from symmetry as a contour consisting of c_2 and its image links the current flowing into the slot (cf. fig. 108). Now since $E_1 = H_2 = \zeta_0 H_1$ and $H_1 = \zeta_0^{-2} E_2 = \zeta_0^{-1} H_2$ we have

$$e_1 = \tfrac{1}{2}\zeta_0 i_2 \tag{10}$$

and
$$e_2 = \tfrac{1}{2}\zeta_0 i_1. \tag{11}$$

Again $e_1/i_1 = Z_1$ and $e_2/i_2 = Z_2$, where Z_1 and Z_2 are respectively the input impedances of the dipole and slot. We obtain therefore the interesting relationship

$$Z_1 Z_2 = \tfrac{1}{4}\zeta_0^2. \tag{12}$$

A relation of the form (12) has been shown by Booker[3] to hold for any two complementary screens, this being only a particular example.

From (12) we deduce, taking $Z_1 = 73\cdot2$ ohms, that for a narrow resonant half-wave slot the resonant resistance Z_r is given by

$$Z_r = 485 \text{ ohms.} \tag{13}$$

The variation of Z_2 with l may, of course, be determined at once from (12), using the value (3) for the characteristic impedance of the equivalent dipole. From the form of (12) we may deduce at once that the variation of resistance and reactance of a slot aerial will be exactly the same as that of the conductance and susceptance of a dipole aerial. We may see the general form of this from fig. 103. In particular, we may note that for a half-wave slot aerial the input resistance has a maximum near the resonant point. The effect of widening the slot will be to reduce the rate of change of reactance just as for a dipole aerial. The 'Q' of equivalent slots and dipoles will, in fact, be equal. As the width of the slot becomes very small the distribution of electric field across it will vary sinusoidally along it, being, of course, zero at the ends. This follows at once from the connexion between the magnetic field and the current in the dipole aerial.

If we assume a sinusoidal distribution of electric field we may evaluate the radiation resistance by a method similar to that given in Chapter 3 for a dipole aerial. This has been done by Bailey[4], who has also considered the characteristic impedance of strip and slot transmission lines. By comparing his formula

$$Z_0 = 120 \log_e \left(\frac{4d}{\omega} \right) \tag{14}$$

for the characteristic impedance of a parallel wire line made of two narrow strips of width ω separated by a distance d with the ordinary formula for a line made up of cylindrical wires of radius a, we get a justification of the usual 'equivalence' formula $\omega = 4a$.

So far we have not considered very closely the physical picture of the operation of the radiating slot. When a field exists across the slot opposite sides will carry opposite charges (fig. 109(a)). These will tend to flow round the ends of the slot, and will build up charges in the opposite sense due to the inductance of the loop round which

they flow. Oscillations will therefore take place in which the charges reverse their sign and oscillatory currents will flow round the ends. These currents may be regarded as the source of radiation, but their distribution is rather difficult to calculate, as they will spread over the metal to a considerable distance from the slot. It is just this spreading which enables a narrow slot to radiate. To a zero order of approximation a narrow slot may be regarded as a pair of $\frac{1}{4}\lambda$ transmission lines in parallel, and to this approximation, of course, could not radiate. Another way to regard the slot is as a linear distribution of electric dipole moments, each elementary dipole being due to the charges on opposite sides of the slot. All these pictures, of course, lead to the same radiation field.

13·3. Measurements on slot aerials

J. L. Putman has measured the input impedance of a slot aerial in a large metal sheet for various widths of slot[5]. The general form of the variation of input resistance and reactance with length is shown in fig. 112. The variation of resistance is just as would be expected, showing a maximum near the resonant point. There is one unexpected feature in the variation of reactance. The variation of resonant length with width of slot seems to be less than for the complementary dipole. This, however, may be due to the method of measurement, as the space between the feeding transmission line could not be varied. The slot was fed by means of an open line consisting of two wires $\frac{3}{8}$ in. in diameter at a fixed distance of 2 in. apart. Thus the amount of metal between the point of attachment of the line and the edge of the slot varied as the width of the slot was changed. Apart from this point the variation of reactance is as expected, being inductive when the length of the slot is less than the resonant value in contrast with a dipole which is capacitive.

Bailey[4] gives measurements which verify that the polar diagram of a slot aerial is the same as that for a half-wave dipole. Much more extensive measurements by Putman, Russell and Walkinshaw[6] have completely verified this and have given the distribution of the field over the surface of the metal sheet containing the slot.

The polar diagram found at a large distance from the slot is exactly as for a $\frac{1}{2}\lambda$ dipole when the size of the metal sheet is large (circle of radius $> 5\lambda$) (fig. 113). For sheets of smaller size, however,

edge effects are noticeable, and the polar diagram shows subsidiary maxima (fig. 114).
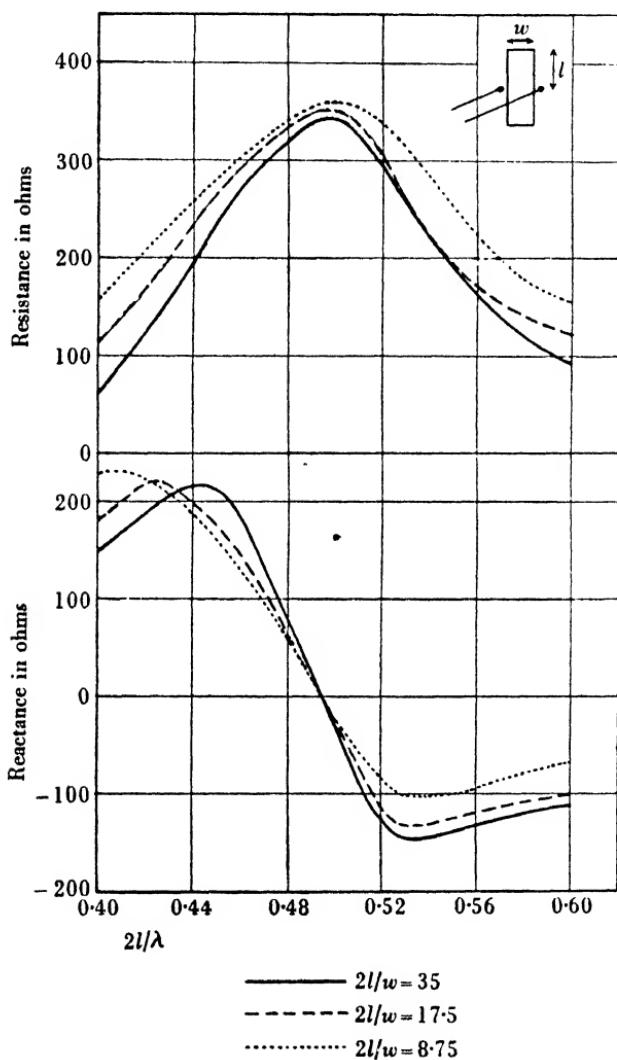
Fig. 112. Input impedance of half-wave slot.

The shape of the equiphase surfaces near the metal screen and in a plane at right angles have also been measured (6). These are found to be confocal ellipses with foci at the ends of the slot.
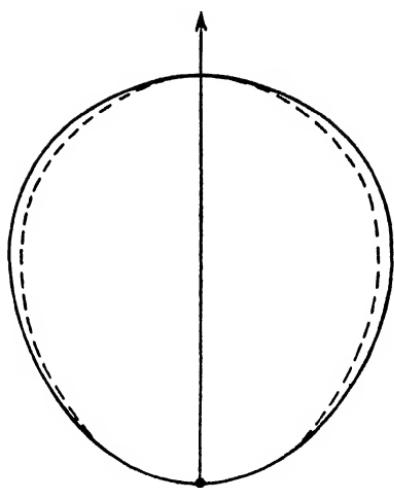
————— Experimental curve (Putman *et al.*)
— — — Theoretical curve for $\frac{1}{2}\lambda$ dipole

Fig. 113. Polar diagram of slot aerial in large metal sheet.
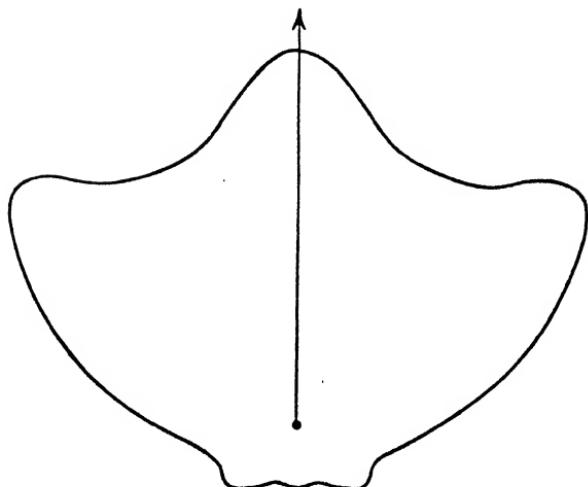
Fig. 114. Polar diagram of slot aerial in circular sheet of radius 1·5λ.

13·4. Cavity-fed slots

It is frequently desirable that a slot should radiate on only one side. This may be achieved by placing a sheet of metal at a distance behind the sheet containing the slot, and parallel to it. It is more usual, however, to enclose the slot entirely on one side of the sheet by means of a metal box. The slot may therefore be regarded as an opening in the cavity, and this gives us another means of feeding the slot. The cavity need not be resonant at the operating frequency, but is often made approximately so in order to store energy. The

(a) Wide-band unipole feed   (b) Dipole feed
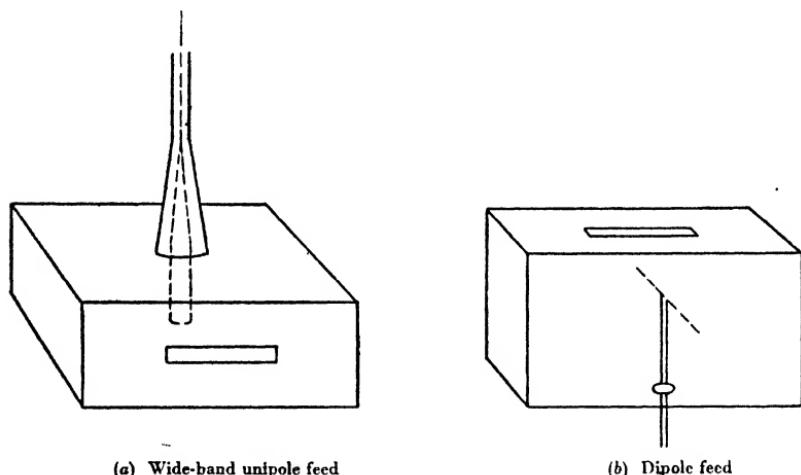
Fig. 115. Methods of feeding a slot by means of a probe in a cavity.

resonance of a cavity with a half-wave slot in it will, however, be quite 'flat', as the slot is an efficient radiator and damps oscillations set up in the cavity. The dimensions of the cavity do not affect the radiation pattern of the dipole, but have a very marked effect on the input impedance. In general, the slot is not fed directly when backed by a cavity, but the cavity is excited by means of a probe as shown in fig. 115. If the cavity is small the impedance is highly reactive and it is difficult to feed.

The band width of a cavity-fed slot may be increased by widening the slot. Frequently a 'dumb-bell' shape is used, the ends being rounded and widened. Also the band width will be increased if a fat probe is used to feed the cavity. A resistive input impedance

of about 50 ohms may be obtained with a nearly resonant cavity fed by a unipole. The length of the unipole is varied to remove the reactive component of the input impedance.

13·5. Applications of slot aerials

Historically the first application of a slot aerial by British workers was to provide a wide-band feed for a large paraboloidal reflector used for an aerial system for 600 Mcyc./sec. radar. Discussion of this will be postponed to Chapter 14. Slot radiators played an important part in the development of linear arrays for microwaves, and this aspect will be fully dealt with in another book in this series dealing with microwave aerials.

13·5·1. *The B.A.B.S. aerial system*

The most important application of resonant slots in the $1\frac{1}{2}$ m. band was to provide a stable aerial for the B.A.B.S. system of instrument landing(7). We have not, so far, discussed the use of shaped metallic reflectors such as paraboloids to produce narrow beams. This is, indeed, beyond the scope of this book. The use of a dipole or unipole in a simple 'corner' is, however, quite extensive, and such aerials have been used to produce overlapping beams for blind landing systems on a wave-length of $1\frac{1}{2}$ m. One arrangement is illustrated in fig. 116. Two unipoles A and B are placed in a metallic 'corner' reflector and energized alternately to produce polar diagrams, as shown in fig. 116(b), whose overlap provides an equi-signal course. The weakness of this system is that independent feeds are required for A and B, and any variations in line characteristics will affect the relative signal strength radiated when A or B is energized and will therefore move the equi-signal course. This is, in fact, found to be a common occurrence.

By having two slots in a cavity fed by a single probe the effect of the variation of feeder characteristics may be entirely eliminated, the switching merely consisting of putting a mechanical short across the centre of one of the slots. The radiation from the short-circuited slot is negligible. The dimensions of the cavity, the thickness of the probe and its position in the cavity are chosen to give a match to the main feeder, when one slot is open and the other closed. It is evident that provided the mechanical switching arrangements remain constant the equality of the two beams and hence the position of their

PLATE III
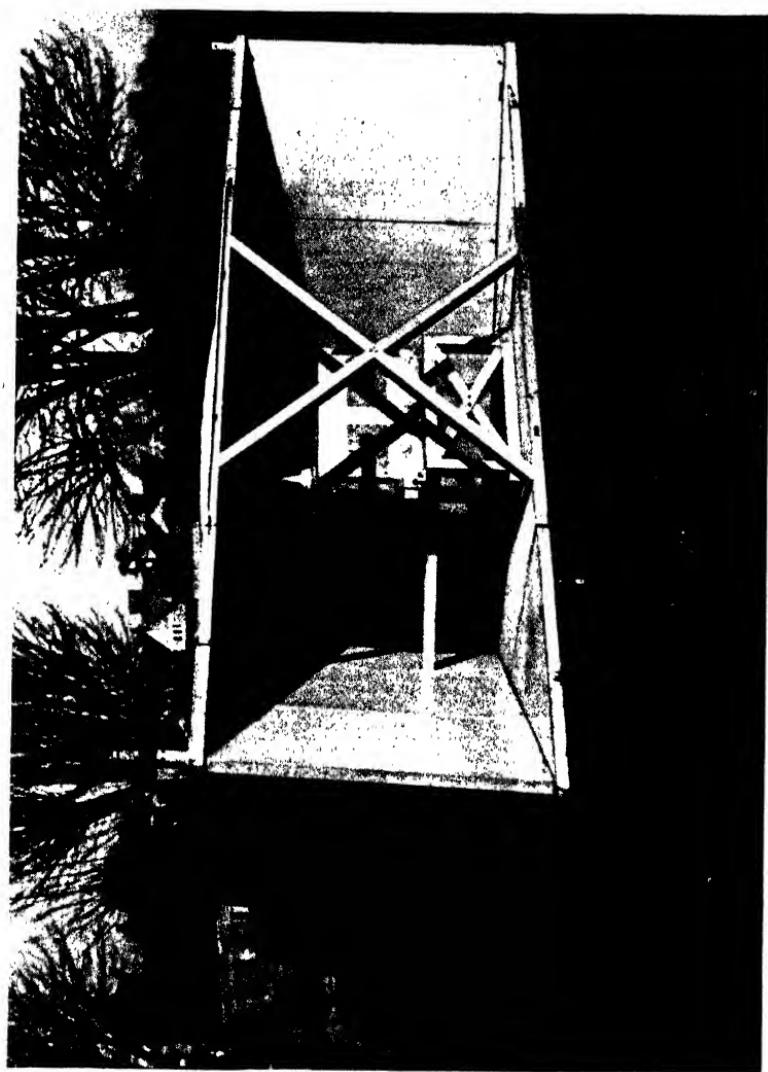
Fig. 117. B.A.B.S. aerial system.

intersection is maintained irrespective of any changes in the probe
and feeder assembly. Such an arrangement has been used for the
aerial system of the B.A.B.S. instrument landing system[8]. The
cavity with its two radiating slots is placed in a 'corner' reflector,
the position of the cavity being chosen so as to produce the desired
polar pattern. The arrangement is shown in the photograph, fig. 117
(Pl. III).

● *A*

● *B*

(*a*)

Equi-signal
line

———— Polar diagram when *B* is excited

– – – – Polar diagram when *A* is excited

(*b*)
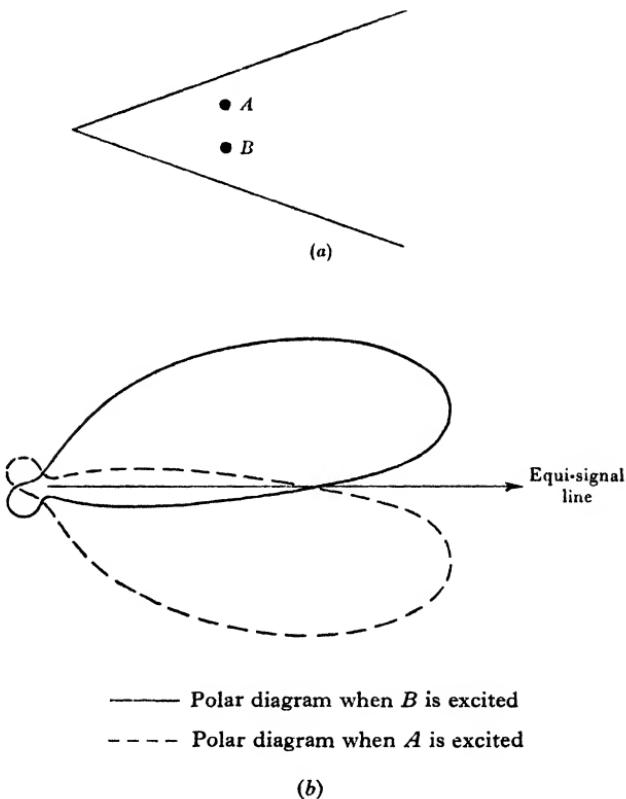
Fig. 116. Use of 'corner' reflector to produce overlapping beams.

13·6. Use of slot aerials in aircraft

Aircraft aerials which do not protrude outside the skin of the
aircraft are obviously desirable for aerodynamic reasons, more
especially for modern very high-speed aircraft. Some investigations
have therefore been carried out on the use of slots in the aircraft
skin as radiators. It is evident that the use of such slot radiators

is limited to very high frequencies, since the increased weight of the structural modifications necessary to accommodate large slots for low frequencies may more than counterbalance the drag of an external aerial if the latter can be properly streamlined. In addition, the positioning of slot aerials on aircraft is just as difficult as for external aerials. The reason for this is that, for very high-speed aircraft, aerodynamic considerations severely limit the regions of the aircraft skin in which slots can be cut. This limitation is quite apart from others which may be imposed by lack of space inside any particular region of the aircraft skin for fitting the necessary cavity for feeding the slot and defining its impedance. A cavity type of feed for slot aerials in aircraft is necessary, since it eliminates the possibility of unwanted changes in the slot impedance which might be introduced by structural modifications or additions inside the skin of the aircraft near the slot. From these considerations it is evident that the design of a slot radiator for a specific application for metre or decimetre wave-lengths on any particular aircraft is best undertaken in the design stage of the aircraft itself, and that the only way of obtaining a system which is flexible in frequency is to choose as high a frequency as possible and make the slot radiator system cover as wide a band as possible. Changes of frequency by retrospective modifications to slot aerials in aircraft would be impracticable.

To obtain uniform transmission or reception with horizontal polarization in the horizontal plane by means of slot radiators in an aircraft it is necessary to use two slots, for example, two vertical slots, one on either side of the fuselage. For vertical polarization the problem is easier. A V-shaped slot or a circular slot cut in the under-surface of the fuselage is adequate. These aerials are still in the experimental stage, but are likely to provide an attractive solution to the problem of providing 'suppressed' aerials for a number of applications.

REFERENCES

(1) Booker, H. G. Slot aerials and their relation to complementary wire aerials (Babinet's Principle). *J. Instn Elect. Engrs*, **93**, IIIA, 620 (1946).
(2) Copson, E. T. An integral-equation method of solving plane diffraction problems. *Proc. Roy. Soc.* A, **186**, 100 (1946).
(3) Booker, H. G. Elements of wave propagation using the impedance concept. *J. Instn Elect. Engrs*, **94**, 171 (1947).

(4) Bailey, C. E. G. Slot feeders and slot aerials. *J. Instn Elect. Engrs*, **93**, IIIA, 615 (1946).

(5) Putman, J. L. Input impedances of centre-fed slot aerials near half-wave resonance. *J. Instn Elect. Engrs*, **95**, 290 (1948).

(6) Putman, J. L., Russell, B. and Walkinshaw, W. Field distributions near a centre-fed half-wave radiating slot. *J. Instn Elect. Engrs*, **95**, 282 (1948).

(7) Smith, R. A. *Radio Aids to Navigation*, Chapter 13. Cambridge University Press (1947).

(8) Wood, K. A. 200 Mc./s. radar interrogator beacon systems. *J. Instn Elect. Engrs*, **93**, IIIA, 481 (1946).

(9) Huxley, L. G. *A Survey of the Principles and Practice of Wave Guides*, p. 283. Cambridge University Press (1947).

Chapter 14

AERIALS FOR DECIMETRE WAVE-LENGTHS

14·0. Introduction

Following the development of radar on $1\frac{1}{2}$ m., the next great step forward was the introduction of microwaves starting with wave-lengths of the order of 10 cm. For the purpose of early warning, however, greater range may be obtained by using a longer wave-length. In order to obtain increased discrimination and better low-angle coverage than is available with a wave-length of $1\frac{1}{2}$ m., developments also took place at an intermediate wave-length near 50 cm. For purposes of radar navigation a band of wave-lengths around 30 cm. was also developed, particularly for a project known as U.N.B. (United Nations Beacons)(1). The technique for this wave band was quite new. It is with aerials for wave-lengths around 50 and 30 cm. that we shall be concerned in this chapter.

Initially, tests were made with scaled-down versions of the broadside and Yagi aerials which had been used at $1\frac{1}{2}$ m. However, the difficulties of correctly phasing the elements and obtaining the correct power distribution among them were very much greater than at $1\frac{1}{2}$ m. Consequently this type of aerial was very soon abandoned in favour of the parabolic reflector, illuminated by a small feed near the focus. Such a system has the following advantages over multi-element arrays. The feeder system is simpler to design and maintain, and since only one fed element is used it is possible to arrange for the system to work over a wide band of frequencies. It is possible to deflect the beam over a small angle simply by moving the fed element a small distance from the focus in the focal plane of the reflector.

Flexibility of transmitter frequency is easy to obtain at these wave-lengths, and this, together with the above considerations, led to the general development of broad-band working. This subject is more closely allied to microwave aerials than to metric aerials. We shall, however, briefly consider some of the dipole and slot arrangements used for feeding the reflectors, as these bring out a number of important principles with which we have dealt in previous chapters.

14·1. Parabolic reflectors

We shall consider only very briefly the properties of parabolic reflectors. These will be dealt with fully in another book of this series on microwave aerials.

The principle of the parabolic reflector is that energy from the radiating source is spread over a large area, namely, the focal plane of the parabola which, because of the fundamental property of the parabola, radiates in phase. Consequently, just as in the case of a broadside array of dipoles, a narrow beam and high gain are produced by this large area of radiating sources (see §3·5). If the primary source were able to produce uniform illumination over the whole area, then the gain would be $4\pi A/\lambda^2$, where A is the aperture area. The beam width, from maximum to zero, is given by $1\cdot22\lambda/D$ radians, D being the diameter. Relative to a rectangular aperture a uniformly illuminated circular aperture gives reduced side-lobes, the first lobe being 13 % of the main lobe instead of 21 %.

In practice, uniform illumination of the aperture is never obtained, for it requires that the radiating source should send four times as much power per unit solid angle towards the edge of the reflector as towards the apex. In fact, most of the sources used give less field strength towards the edge of the reflector and, of course, a lot of energy would be wasted if they did otherwise. This results in a broadening of the main beam, diminution of gain and reduction of side-lobes relative to a uniformly illuminated aperture.

14·2. Wide-band feeds for 50 cm. aerials

For a number of applications wide-band dipole feeds are used for supplying the primary energy at the focus of the paraboloid. These are simply fat half-wave cylindrical dipoles with stub compensation (§ 12·7). A typical example is shown in fig. 118. Westcott and Goward have given a comprehensive account of the various methods of wide-band matching used for obtaining the best results from the use of these dipoles[2].

14·2·1. *Slot radiators for 50 cm. wave-lengths*

For early-warning radar equipment operating on a wave-length near 50 cm. it was desired to use vertical polarization, as echoes from the sea are less with vertical polarization at this wave-length. It was required to feed a wide section of a paraboloid from a single

source, and if a horizontal dipole is used the edges of the paraboloid are insufficiently illuminated owing to the $\cos\theta$ pattern of the dipole. A source of horizontally polarized radiation is therefore required which gives an almost uniform polar diagram over 180° in the horizontal plane. It was for this purpose that the first slot aerial was developed mainly by Westcott[2] and Booker[3].

For use as a feed for a parabolic reflector it is, of course, desirable that radiation from the source should only be directed *towards* the reflector. The slot is therefore backed by a cavity. In this case the
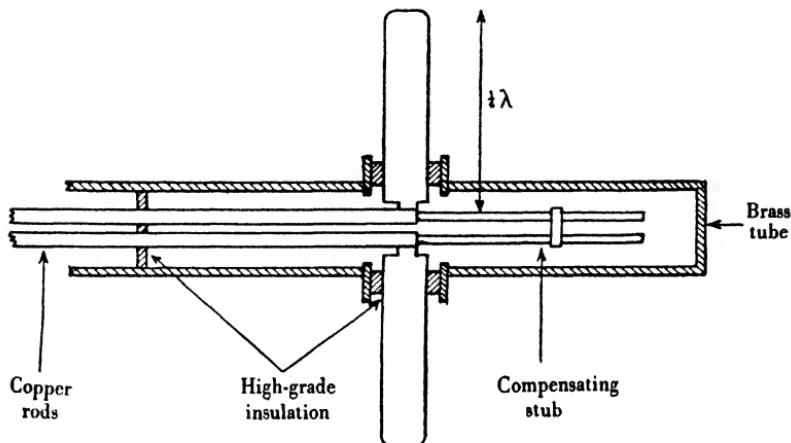
Fig. 118. Wide-band dipole.

cavity is fairly small, so as to be easily supported at the focus of the reflector. For this aerial the cavity takes the form of a cylindrical can with a vertical slot cut in it as shown in fig. 119.

14·3. Wide-band dipoles and unipoles for wave-lengths near 30 cm.

A considerable variety of wide-band dipoles and unipoles were developed for use at wave-lengths around 30 cm. by the Combined Research Group* in the U.S.A. for use both as feeds for reflectors and for 'all-round' aircraft aerials. These again were based on the principle of using fat dipoles with reactance compensation. Use was widely made of conical feeds (see Chapter 12) and the conical trans-

* This group was specially formed to study wave-lengths near 30 cm. It consisted of a combined team of American and British research workers.

former was often the source of reactance compensation.* Typical examples of these aerials are shown in figs. 120–122. Fig. 120 shows two $\frac{1}{4}\lambda$ wide-band radiators. That in 120(b) consists of a section raised above an earthed plane. This has the effect of raising the input
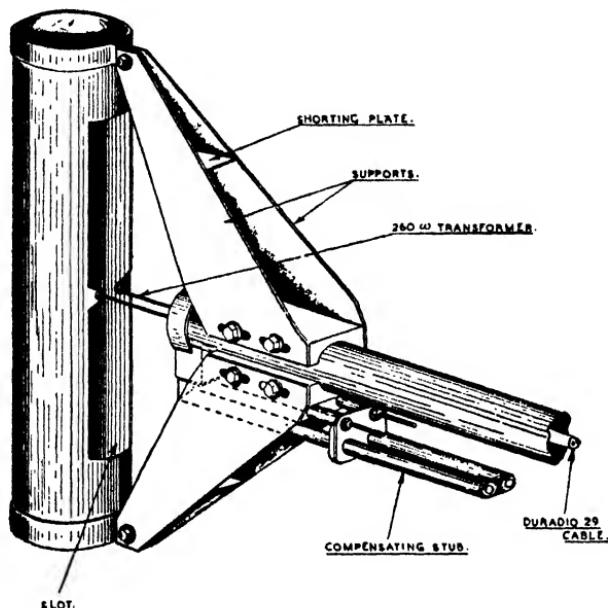
(a) Practical construction
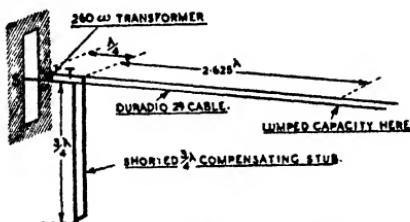
(b) Electrical details

Fig. 119. Matched slot and feeder system for 50 cm. radar aerial.

impedance and facilitates wide-band compensation. Fig. 121 illustrates the use of a conical radiator to give a half-wave wide-band unipole (cf. fig. 101). Unipoles are mainly used for this work as the

---

* The author is indebted to Mr V. H. Rumsey and Mr R. Aspinall for advance information on these aerials, prior to publication.

conical feed is a most useful device for matching purposes, and for reactance compensation.

An interesting example of a 'suppressed' aerial is shown in fig. 122. This is intended for mounting flush with the metallic skin of an aircraft. One way of looking at this aerial is to regard it as a circular slot radiator fed by a coaxial line.
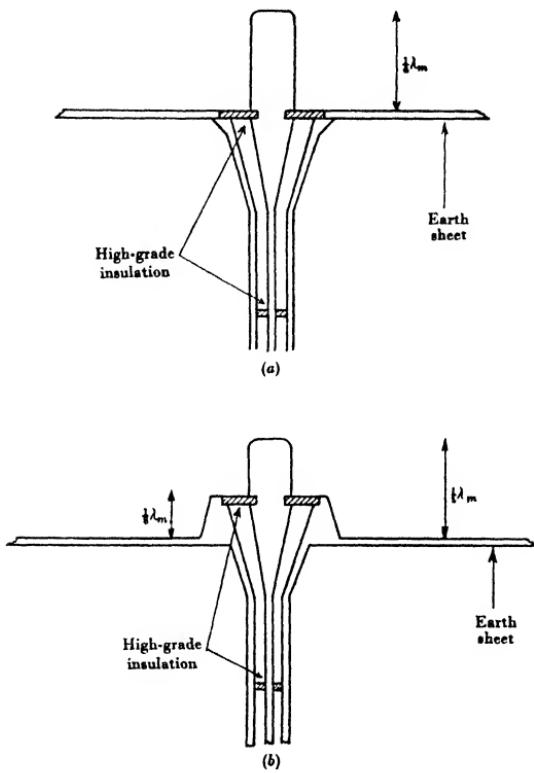
Fig. 120. Unbalanced quarter-wave radiators for 30 cm. wave-length.

14·4. Wide-band slot aerials for 30 cm.

Slot aerials have also been used for wave-lengths in the 30 cm. band. The slot switching technique as described in connexion with the B.A.B.S. aerial system in Chapter 13 was used to provide overlapping patterns for direction finding from aircraft. In fact, it was early work on an aerial for this purpose by V. H. Rumsey that suggested the use of switched slots for the B.A.B.S. aerial. An aerial for direction finding in aircraft is shown in fig. 123.
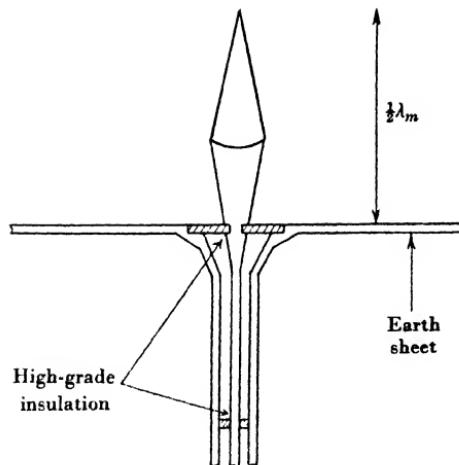
Fig. 121. Conical half-wave unipole.
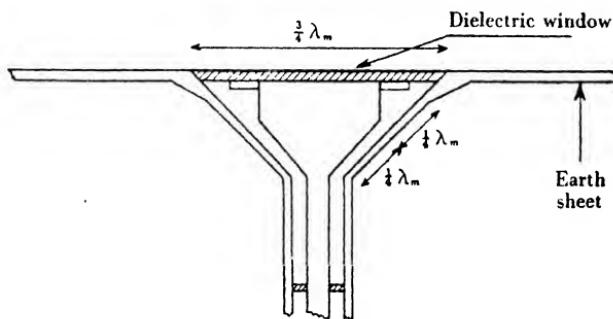
Fig. 122. Submerged quarter-wave radiator, or circular slot radiator.
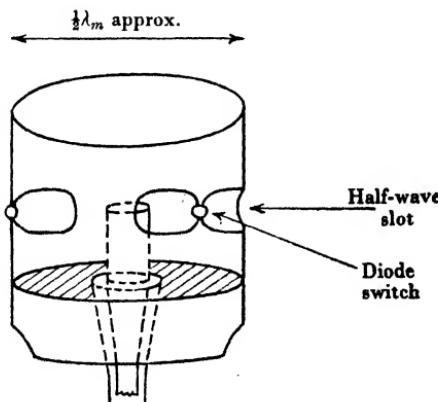
Fig. 123. Double slot aerial for direction finding in aircraft.
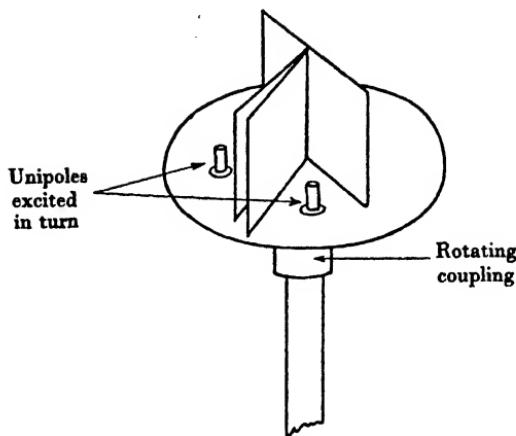
Fig. 124. Aerial for direction finding in aircraft.
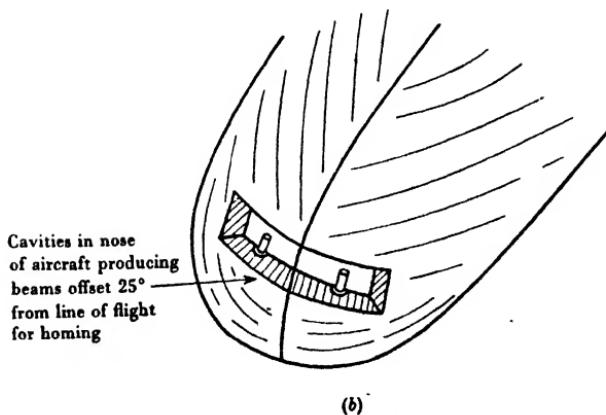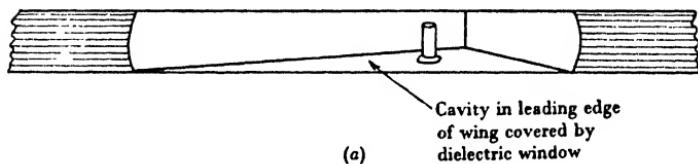
Fig. 125. Aerials mounted in the leading edge of a wing
and in the nose of an aircraft.

14·5. Aerial systems for 30 cm. applications

A number of aerial systems have been developed for 30 cm. applications. The small wave-length makes the use of shaped cavities and reflectors for producing directional patterns relatively easy. An alternative arrangement to that shown in fig. 123 for producing overlapping patterns for direction finding is shown in fig. 124. This is housed in a streamlined rotatable 'egg'. It is interesting to note that at this wave-length, direction finding, as distinct from homing, is possible. Only at very short and very long wave-lengths can interference from the aircraft structure be reduced sufficiently to obtain worth-while bearings.

Another consequence of the short wave-length is that aerials may be mounted in the leading edge of the wings or in the nose of an aircraft as shown in fig. 125. Diffraction round the fuselage and wings at these short wave-lengths is small, and if complete 'all-round' cover is required more than one aerial must in general be used.

REFERENCES

(1) Smith, R. A. *Radar Aids to Navigation*, Chapter 3, Cambridge University Press (1947).
(2) Westcott, C. H. and Goward, F. K. Design of wide-band aerial elements for 500–600 Mc./s. ground radar. *J. Instn Elect. Engrs* (in press).
(3) Booker, H. G. Slot aerials and their relation to complementary wire aerials (Babinet's principle). *J. Instn Elect. Engrs*, **93**, III A, 620 (1946).

Chapter 15

NOISE IN AERIALS

15·0. Limit of sensitivity set by noise

The limiting sensitivity of a receiver as regards its ability to pick up and display signals is determined by the fluctuation 'noise' in its output. This 'noise' consists of random electrical fluctuations and is generally separated into three parts:

(1) That picked up by the aerial and amplified.

(2) That generated in the first tuned circuit of the receiver, and amplified.

(3) That produced by the rest of the receiver.

It is well known that a resistance R at absolute temperature T generates in a band width B a random e.m.f. (Johnson noise) whose mean square value $\overline{e^2}$ is given by[1]

$$\overline{e^2} = 4kTBR, \qquad (1)$$

where k is Boltzmann's constant.

Now suppose we have a passive load with resistive component R connected across the input of a receiver. Let it be at the same temperature as the receiver and let R be equal to the input resistance of the receiver. If matched to the receiver the load will feed to it a power P_0 given by

$$P_0 = kTB. \qquad (2)$$

If the receiver were perfect and itself generated no noise this would be the limiting power input for which signal would be equal to noise. In fact the receiver does generate noise. This is made up of Johnson noise from the input circuit and 'shot' noise in its valves. The power input for which signal will be equal to noise will then be given by

$$P_0 = NkTB, \qquad (3)$$

where N is a factor known as the noise factor of the receiver. It is usually expressed in decibels.

15·1. Aerial noise

In general, when an aerial is matched to a receiver the noise power fed from it to the input will not be equal to kTB. It is common,

however, to ascribe to the aerial an effective temperature T_e defined so that the noise fed to a matched receiver is kT_eB. The matched condition permits maximum transfer of power to the receiver, but this may not be the condition for maximum signal-to-noise ratio at the input. A slight improvement may be obtained with a mismatched condition. This is a somewhat difficult question and belongs to the study of the design of input circuits for receivers. The matched condition will, however, give a very good indication of the minimum signal detectable under given conditions.

Suppose the aerial were to receive its signals from regions of space all at temperature T', then it is very easy to show(2) that its effective temperature is just T' and its effective resistance for noise production is just equal to its radiation resistance provided ohmic loss in the aerial may be neglected. Consider such an aerial, whose radiation resistance is R, placed in a uniform temperature enclosure whose temperature is T'. Let it be matched to a resistance R whose temperature is also T'. The system will then be in thermal equilibrium. The power fed from the resistance to the aerial will, from (2), be $kT'B$, and this must be equal to that fed to the resistance from the aerial. This is so for any temperature T'. We may therefore regard the aerial as equivalent to a resistance R at temperature T' and producing noise in a band width B with mean square e.m.f. given by

$$\overline{e^2} = 4kT'BR. \tag{4}$$

If now the aerial is matched to a resistance at temperature T, it will either feed power to or draw power from this resistance according as $T' > T$ or $T' < T$. If the actual temperature of the aerial is equal to T and it has ohmic resistance R_0 and radiation resistance R_r, the power fed to the load R will be

$$\frac{kB}{R}[T'R_r + TR_0], \tag{5}$$

if $R = R_0 + R_r$, since the mean powers from the ohmic and radiation resistance are additive, both being due to random fluctuations. The effective temperature T_e of the aerial is then given by

$$RT_e = T'R_r + TR_0. \tag{6}$$

For resonant aerials R_0 is usually negligible, but not for terminated aerials.

In general, the noise received by an aerial will not originate in places at a uniform temperature, but it is not difficult to generalize the expression for the effective temperature of the aerial[2]. Let $G(\theta, \phi)$ be the gain of the aerial and $T(\theta, \phi)$ be the effective temperature of the region in a direction (θ, ϕ). Then, if we neglect ohmic loss,

$$T_e = \frac{1}{4\pi} \iint T(\theta, \phi)\, G(\theta, \phi) \sin\theta\, d\theta\, d\phi, \tag{7}$$

which follows at once from the fact that powers received from different directions are additive due to the random nature of the noise. In particular, if radiation is only received over a small solid angle ω from an object at temperature T, the aerial being directed towards the object and its gain being practically constant over ω, then

$$T_e = T\omega G/4\pi. \tag{8}$$

If some of the radiation from the aerial is directed towards the ground and an appreciable fraction is absorbed we must take the temperature of the ground into account. For small angles of incidence we have seen that the ground acts as a perfect reflector for metre wave-lengths, but this is no longer true for large angles of incidence which must be taken into account when we consider the noise picked up by the aerial. This effect has been treated in some detail by Burgess[2].

The use of an effective temperature for an aerial may still be retained, even when the radiation received is not thermal radiation, so long as it consists of random fluctuations. The effective temperature of any region of space may be defined as that temperature which would produce the same density of thermal radiation at the frequency of observation. The effective temperature of the aerial may then be defined by (7).

We must now consider the various sources of noise which cause such fluctuations in aerials. We must then determine their relative importance as compared with receiver noise in setting the limit to the sensitivity of a receiving arrangement.

15·2. Sources of aerial noise

There is still some doubt as to where all the noise received in aerials originates. Most of the sources are, however, now known, though there is still doubt in some cases as to the mechanism whereby

the noise is produced. It is fairly certain that most of the incoming noise on metre wave-lengths is not due to black-body radiation, and the effective temperature derived from the intensity of the radiation does not represent the equilibrium temperature of the places of origin.

The known sources of aerial noise are as follows:

(1) Thermal agitation noise due to loss resistance in the aerial.

(2) Thermal radiation from nearby objects such as the ground or from hot objects such as the sun.

(3) Abnormal radiations from the sun.

(4) Radiations having their origin in the Milky Way, known as cosmic or galactic noise.

(5) Noise due to atmospheric disturbances such as lightning flashes.

(6) Noise generated in the ionosphere.

(7) Man-made noise.

The noise due to ohmic loss in the aerial is, as we have seen, negligible for most aerials used at metre wave-lengths, except for terminated aerials. Man-made noise varies greatly with locality, but except in regions where a great deal of electrical machinery is operating is not in general an important factor in limiting sensitivity. In aircraft, however, ignition noise from the engines may be very troublesome unless these are carefully screened. Careful bonding of various parts of the aircraft is also necessary if excessive noise is to be avoided.

For wave-lengths such that long-distance transmission via the ionosphere is possible noise due to 'static', mainly caused by thunderstorms, is the most important source. This noise completely swamps that generated by any good receiver and sets the limit to sensitivity provided a reasonable aerial is used. This applies mainly to wave-lengths greater than 20 m. As the wave-length is reduced to 10 m., most of this noise disappears and noise from extra-terrestrial sources is predominant. Between 20 and 10 m. the noise is probably a mixture of atmospheric and cosmic noise, though it is very difficult to separate the two. It would be expected that, if atmospheric noise were of much consequence, the noise level would show a marked daily variation and a preponderance of noise at small angles of elevation. Attempts to discover this effect using C.H.

radar receivers and aerials failed to show it.* It was found that at 10 m. the noise power was remarkably steady and 15–20 times the equivalent first circuit noise of the receivers used. Although more careful recent experiments have shown a random fluctuation in this noise level and a diurnal variation of small amplitude, even when a non-directional aerial is used, the conclusion is that not much of the noise received on an aerial at 10 m. is due to atmospherics.

It was thought at one time that a considerable part of this noise might be due to fluctuations of electric charges in the ionosphere. Although there may, in fact, be a residual part of the noise due to this cause, it is now known that most of the noise at wave-lengths of the order of 10 m. and less comes from well-localized regions of the Galaxy. We must now consider in some detail this important source of noise which in general sets a limit to sensitivity for wave-lengths between about 15 and 5 m. It may also limit sensitivity in certain well-defined directions for shorter wave-lengths if highly directive aerials are used. In general, however, for wave-lengths shorter than 5 m. sensitivity is limited by receiver noise.

15·3. Cosmic or galactic noise

The discovery that aerial noise in the metre band is mainly of galactic origin is due to Jansky[3], who noted that at a wave-length of 15 m. the noise came predominantly from the direction of the Milky Way. The noise received on a half-wave dipole will therefore be different from that received by a highly directional aerial. When the latter is pointed towards the Milky Way the noise will be considerably greater than for a half-wave dipole. Exact measurements of effective temperature of a half-wave aerial for various frequencies are somewhat scanty. The measurements using C.H. radar receivers gave an input power of the order of 2×10^{-12} W. for a 1 Mcyc./sec. band width at 12 m. and 5×10^{-13} W. at 6 m. These values are now thought to be about 8 db. too high. If we reduce them by this amount they fit well with other measurements[3, 4]. These indicate that we may take as a mean temperature for outer space a value of about 3000° K. at 6 m., and there is some evidence that the mean temperature varies approximately as λ^2.

The effective temperature of the Milky Way is considerably

* Unpublished experiments by Dr B. B. Kinsey (1941).

higher and varies with galactic latitude. The highest effective temperature appears to be of the order of 10,000° K. at 60 Mcyc./sec. and to vary (4) as $\lambda^{2 \cdot 7}$. The mean noise level picked up on a half-wave dipole varies during the day by only a few db., but that picked up on a directional aerial may rise as much as 7–8 db. above the mean level.

Table 15 shows the relation between power input in watts and effective temperature for a band width of 1 Mcyc./sec. It also shows the value of the noise factor for a receiver at 300° K. required to make the aerial noise equal to the receiver noise.

Table 15

Effective aerial temperature (° K.)	Aerial noise in watts for 1 Mcyc./sec. band width	Noise factor of receiver to make set noise equal to aerial noise (db.)
30,000	4×10^{-13}	23·0
10,000	$1 \cdot 3 \times 10^{-13}$	18·2
3,000	4×10^{-14}	13·0
1,000	$1 \cdot 3 \times 10^{-14}$	8·2
300	4×10^{-15}	3·0

If the effective aerial temperature is known, table 15 may also be used to find the power ratio of aerial noise to set noise when the noise factor of the receiver is known. This ratio in db. will simply be the difference between the figure in the last column of table 15 and the noise factor expressed in db. For example, if the effective aerial temperature at 12 m. is 12,000° K., corresponding to a noise power of $1 \cdot 5 \times 10^{-13}$ W. for a 1 Mcyc./sec. band width, a receiver with noise factor 19 db. would have receiver noise equal to aerial noise. A reasonable value for the noise factor of a good receiver at 12 m. is 6 db., so that the aerial noise would exceed set noise by 13 db. At 3 m., on the other hand, assuming the λ^2 law, we should expect an effective aerial temperature of only 750° K. for a non-directional aerial, so that a receiver having a noise factor of 7 db. would be required to give set noise equal to aerial noise. The best receivers in this band have in fact noise factors of the order of 7 db. At $1\frac{1}{2}$ m., still assuming the λ^2 law, the aerial noise in a non-directional aerial would correspond to a temperature of 190° K. and would be quite negligible in an ordinary receiver. This, however, does not hold

for highly directional aerials, since, as we have seen, certain regions of the Galaxy have effective temperatures much higher than the mean value.

15·4. Solar noise

The most important source of thermal noise is the sun. This noise, however, can only be observed under normal conditions by using very highly directive aerials. Let us consider what radiation we should expect from the sun regarded as a black body at 6000° K. From equations (3) and (8) we see that thermal noise will be comparable with receiver noise if

$$NT_0 = \omega G T_s/4\pi, \qquad (9)$$

where T_0 is the receiver temperature, T_s the solar temperature, ω the solid angle subtended by the sun and G the aerial gain. For the sun $\omega = 7 \times 10^{-5}$, so we have for $T_0 = 300°$ K.,

$$\frac{4\pi}{G} = 0.8\omega, \qquad (10)$$

if we take $N = 25$ (14 db.).

Thus we see that we shall require a gain of the order of 2×10^5, or as we may note from (10) an aerial which produces a beam with solid angle of the order of that subtended by the sun.

By using special methods and 'chopping' the solar radiation, Dicke has shown that it may be observed under normal conditions[5]. However, in general, it will be seen that thermal radiation from the sun is not of practical importance. But there are times when intense radiation from the sun is observed on metre wave-lengths. These abnormal conditions are associated with intense sunspot activity and do not correspond to thermal radiation[6]. They would require a temperature of the order of 10^{10} deg. K.

15·4·1. *Abnormal solar radiation*

A good deal of evidence has now been collected on the abnormal solar radiation. At times of great sunspot activity the intensity of radiation at wave-lengths less than those absorbed by the ionosphere (about 15 m.) and down to about 50 cm. is greatly increased. The amount observed on several occasions would correspond to a solar temperature of 10^{10} deg. K. if it were black-body radiation[7]. The radiation has now been shown by Appleton and Hey[8] and by Ryle

and Vonberg[9] to be circularly polarized, and a measurement of the angular extent of the places of origin by the latter authors has shown that it almost certainly arises from large sunspots. The maximum value of the radiation density observed by Lovell and Banwell[7] was about $8 \cdot 6 \times 10^{-17}$ W./sq.m./c.p.s. at a frequency of 72 Mcyc./sec., corresponding to $1 \cdot 3 \times 10^8$ times normal thermal radiation, but the average value during the period of intense activity which they observed was about 5×10^{-20} W./sq.m./c.p.s. Such intense radiation has been observed at frequencies ranging from about 30 to 200 Mcyc./sec., and a reduced effect appears even at 500 Mcyc./sec. On centimetric wave-lengths there appears to be only a relatively small increase. These conditions appear only at times of intense sunspot activity and last for a few days. Only then do they have an appreciable effect on radio reception in the metre band.

If we assume a value of 10^{-19} W./sq.m./c.p.s. as typical for the power flux of solar noise at 12 m. during abnormal conditions, we may calculate the power picked up by a half-wave dipole. The effective absorbing area is $\frac{1}{2}\lambda^2$ (§ 3·5) or 18 sq.m., so that the power fed to a receiver with 1 Mcyc./sec. band width will be of the order of 1×10^{-12} W., allowing a factor 2 for polarization. This is about three times the value now generally accepted for the mean cosmic noise level and should be readily observed even using a half-wave dipole as an aerial.

15·5. Effect of aerial noise on effective noise factor of a receiver

Suppose we have a receiver whose noise factor is N. Then the noise generated by the receiver plus input load, if this is a matched resistance at the same temperature, is equivalent to an input power $NkTB$ to a perfect receiver, as we have seen. The equivalent input noise due to the receiver alone will therefore be $(N-1)kTB$. If, however, we have an aerial noise corresponding to an effective aerial temperature T', the total noise power input will be

$$kB[T' + (N-1)T]. \tag{11}$$

If the receiver were perfect the noise power would be

$$kT'B. \tag{12}$$

We may therefore define an effective noise factor N_e in the presence of aerial noise as the ratio of the noise powers given by (11) and (12), i.e.

$$N_e = 1 + (N-1)\,T/T'. \tag{13}$$

Clearly if $T' \gg T$, $N_e \to 1$ and the receiver behaves as a perfect one. The efficiency of the aerial may then be dropped without change of signal-to-noise ratio till $(N-1)\,T/T'$ becomes comparable with unity without affecting signal-to-noise ratio. The fall off in efficiency will, of course, correspond to a drop in T'.

Summarizing these conclusions we may say that for efficient resonant aerials sensitivity is determined by aerial noise for wave-lengths greater than 5 m. Using the very best grounded grid valves for the first stage of the receiver sensitivity is affected by aerial noise in certain directions down to wave-lengths of the order of 3 m., but is generally determined by receiver noise. However, under conditions of abnormal sunspot activity the sensitivity may be greatly reduced at wave-lengths greater than a metre, particularly when directional aerials are used looking towards the sun. For shorter wave-lengths sensitivity is entirely determined by receiver noise.

REFERENCES

(1) Johnson, J. B. Thermal agitation of electricity in conductors. *Phys. Rev.* **32**, 97 (1928).
(2) Burgess, R. E. Fluctuation noise in receiving aerials. *Proc. Phys. Soc.* **53**, 293 (1941); **58**, 315 (1946).
(3) Jansky, K. G. Directional studies of atmospherics at high frequencies. *Proc. Inst. Radio Engrs, N.Y.*, **20**, 1920 (1932); **21**, 1387 (1933).
(4) Moxon, L. A. Variation of cosmic radiation with frequency. *Nature, Lond.*, **158**, 758 (1946).
(5) Dicke, R. H. Measurement of thermal radiation at microwave frequencies. *Rev. Sci. Instrum.* **17**, 268 (1946).
(6) Appleton, E. V. Departure of long-wave solar radiation from black-body intensity. *Nature, Lond.*, **156**, 534 (1945).
(7) Lovell, A. C. B. and Banwell, C. J. Abnormal solar radiation. *Nature, Lond.*, **158**, 517 (1946).
(8) Appleton, E. V. and Hey, J. S. Circular polarization of solar radio noise. *Nature, Lond.*, **158**, 339 (1946).
(9) Ryle, M. and Vonberg, D. D. Solar radiation on 175 Mc./s. *Nature, Lond.*, **158**, 339 (1946).

INDEX

DATE OF ISSUE

This book must be returned within 3, 7, 14 days of its issue. A fine of ONE ANNA per day will be charged if the book is overdue.